

Universidad de Santiago de Compostela

Escola Politécnica Superior

Departamento de Producción Vegetal y Proyectos de Ingeniería



ÍNDICES FENOTÍPICOS DE SELECCIÓN EN CAÑA DE AZÚCAR

Memoria presentada para optar al Grado de Doctor por la
Universidad de Santiago de Compostela por el Ingeniero Don Fredy Uber
Rosales Longo

Directores: Dr. D. Santiago Pereira Lorenzo

Dr. D. José Luis Quemé de León

Dr. D. Mario Melgar Morales.

Lugo, 2017





DECLARACIÓN DEL AUTOR DE LA TESIS

ÍNDICES FENOTÍPICOS DE SELECCIÓN EN CAÑA DE AZÚCAR

D. Fredy Uber Rosales Longo

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento, y declaro que:

- 1) *La tesis abarca los resultados de la elaboración de mi trabajo.*
- 2) *En su caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
- 3) *La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.*
- 4) *Confirmando que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.*

En Lugo, de septiembre de 2017

Fdo.....



AUTORIZACIÓN DEL DIRECTOR / TUTOR DE LA TESIS

ÍNDICES FENOTÍPICOS DE SELECCIÓN EN CAÑA DE AZÚCAR

D. Santiago Pereira Lorenzo, Dr. Ingeniero Agrónomo, Catedrático de la Universidad de Santiago de Compostela (USC) del Área de Producción Vegetal y Proyectos de Ingeniería.

D. José Luis Quemé de León, Dr. Ingeniero Agrónomo, Profesor horario de la Universidad de San Carlos de Guatemala (USAC) del área de Estadística.

D. Mario Francisco Melgar Morales, PhD, Director General de CENGICAÑA.

INFORMA/N:

*Que la presente tesis, corresponde con el trabajo realizado por D. **Fredy Uber Rosales Longo**, bajo mi dirección, y autorizo su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como director de ésta no incurre en las causas de abstención establecidas en Ley 40/2015.*

En Lugo., de septiembre de 2017





DEDICATORIA

A Guatemala, la tierna indiana de los pies descalzos.

A mi esposa Ruth Emilsa Rodríguez Gallardo y a mis hijos Juan Carlos y María Emilsa.

A mi madre, Maurilia Longo y mi hermano Juan Carlos Rosales Longo.



AGRADECIMIENTOS

Al Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICAÑA) y su Junta Directiva, por el apoyo en la autorización y la provisión de los recursos necesarios para la ejecución de este trabajo.

Al Dr. D. Santiago Pereira Lorenzo, por el constante apoyo en la realización de este trabajo, por las puntuales recomendaciones y por el aporte de su valiosísimo tiempo; pero sobre todo, por la amistad.

Al Dr. D. José Luis Quemé por su apoyo, asesoría, revisiones y amistad.

Al Dr. D. Mario Melgar Morales, Director General de CENGICAÑA, por el permanente soporte, interés y cooperación, para que el programa de Doctorado Internacional en Agricultura y Medioambiente encontrara un buen camino y rindiera frutos.

Al Dr. D. Agustín Merino, coordinador del programa de Doctorado en Agricultura y Medioambiente (DIAM), por el entusiasmo y trabajo puestos para que me uniera al DIAM y llevara feliz término el programa.

Al personal de campo de los ingenios guatemaltecos Pantaleón y Magdalena que contribuyó formidablemente en la instalación y cuidados agronómicos de los ensayos que permitieron la recolección de los datos correspondientes, en particular al Ing. Vinicio Paz Chávez y al Ing. Luis Lima.

Al personal del laboratorio Agroindustrial de CENGICAÑA, quienes estuvieron siempre prontos en los análisis de jugos y la entrega oportuna de los resultados.

Al Lic. Estuardo Catalán de la Biblioteca Pedro Escaler de CENGICAÑA, quien estuvo presto a la búsqueda y adquisición de mucha de la literatura que se utilizó en la redacción de esta memoria de tesis.

A la Dra. Dña Ana María Ramos Cabrer de la USC, por su ayuda en la revisión, elaboración y maquetación del documento.

Para todos aquellos que de una u otra forma ayudaron en la realización de este trabajo.



RESUMO

En cana de azucre, con grandes poboacións suxeitas de selección, o grao de precisión da selección pode verse menoscabado pola dificultade de realizar a valoración simultánea dun elevado número de variables en cada un dos xenotipos en avaliación. Por tanto, a redución do número de variables nunha ou poucas dimensións (índices) permitiría avaliar e clasificar aos xenotipos dunha maneira eficiente, o que debería conducir a un proceso de selección con mellores ganancias xenéticas. Para iso expuxéronse os obxectivos específicos seguintes: a) avaliar nove metodoloxías para o cómputo de índices de desempeño ou selección en xenotipos de cana de azucre do Estado III de Selección do Programa de Variedades de CENGICANÁ en Guatemala; b) determinar o grao de axuste dos modelos dos índices; e c) identificar as variables que mellor explican o desempeño dos xenotipos de cana de azucre. Para este efecto, establecéronse catro experimentos con 513 xenotipos en dúas localidades diferentes, e foron avaliados en dous anos consecutivos mediante 14 variables fenotípicas (a partir de 58 medicións). En total obtivéronse sete conxuntos de datos. A información analizouse mediante análises estatísticas descritivas unidimensionais, probas de “t”, análise de varianza (ANOVA), análise de modelos lineais mixtos, análises de regresión lineal, análise multidimensional de varianza (MANOVA), análise de compoñentes principais (ACP) e regresións Logit. Estimáronse os scores de nove índices de selección (desempeño): Smith, Kempthorne e Nordskog, ESIM, RESIM, CP1, CP2, CP3, TAH e BLUP. As diferenzas en desempeño dos xenotipos para TAH foron altamente significativas, coa compoñente xenética como o factor máis determinante para a diferenciación, aínda que a interacción *xenotipo* × *ambiente* indicou a necesidade de ensaios comparativos en diversos ambientes. O pobre desenvolvemento en biomasa en dous dos ensaios non permitiu discriminar entre xenotipos en ningún dos índices aplicados. O axuste dos índices de selección Smith, KN, ESIM, RESIM, e CPs non presentaron un patrón xeneralizado, polo que o axuste destes índices debe ser priorizados para cada conxunto de datos. Non se atopou unha función que ponderase apropiadamente todas as 14 variables en estudo. Os ACPs, as análises de varianza e as funcións lineais dos índices de desempeño indicaron que as variables con maior peso na variación foron: **brixprom_1**, **pol_6**, **brixjugo_7**, **purity_8**, **peso5talks_2**, **diam_11**, **alt_10**, **fiber_14**, **flower_12**, **pith_13** e **oqueinter_4**. A variable **flower_12** foi de interese unicamente na diferenciación xenética de variedades florecedoras. A selección

de xenotipos de cana de azucre segundo as regresións **Logit** foi función dos valores xenéticos determinados unicamente polo TAH e BLUP. BLUP optimizou os procedementos de selección de xenotipos e pode ser aplicado de forma rutineira nos procesos de avaliación e selección de xenotipos en condicións de cultivo en Guatemala. Por todo o anterior, o procedemento estatístico máis apropiado para diferenciar xenotipos de cana de azucre foi o Mellor Estimador Lineal Inssegado (BLUP) a partir da variable TAH. A utilización desta metodoloxía permitiu a selección de 25 xenotipos de cana de azucre cun incremento media do 22% no TAH (5.13 t) respecto das testemuñas utilizadas.



RESUMEN

En caña de azúcar, con grandes poblaciones sujetas de selección, el grado de precisión de la selección puede verse menoscabado por la dificultad de realizar la valoración simultánea de un elevado número de variables en cada uno de los genotipos en evaluación. Por tanto, la reducción del número de variables en una o pocas dimensiones (índices) permitiría evaluar y clasificar a los genotipos de una manera eficiente, lo que debería conducir a un proceso de selección con mejores ganancias genéticas. Para ello se plantearon los objetivos específicos siguientes: a) evaluar nueve metodologías para el cómputo de índices de desempeño o selección en genotipos de caña de azúcar del proyecto “*Estado III de Selección*” del Programa de Variedades de CENGICANÑA en Guatemala; b) determinar el grado de ajuste de los modelos de los índices; y c) identificar las variables que mejor explican el desempeño de los genotipos de caña de azúcar. Para este efecto, se establecieron cuatro experimentos con 513 genotipos en dos localidades diferentes, y fueron evaluados en dos años consecutivos mediante 14 variables fenotípicas (a partir de 58 mediciones). En total se obtuvieron siete conjuntos de datos. La información se analizó mediante análisis estadísticos descriptivos unidimensionales, pruebas de “t”, análisis de varianza (ANOVA), análisis de modelos lineales mixtos, análisis de regresión lineal, análisis multidimensional de varianza (MANOVA), análisis de componentes principales (ACP) y regresiones **Logit**. Se estimaron los *scores* de nueve índices de selección (desempeño): Smith, Kempthorne y Nordskog, ESIM, RESIM, CP1, CP2, CP3, TAH y BLUP. Las diferencias en desempeño de los genotipos para TAH fueron altamente significativas, con la componente genética como el factor más determinante para la diferenciación, aunque la interacción *genotipo* × *ambiente* indicó la necesidad de ensayos comparativos en diversos ambientes. El pobre desarrollo en biomasa en dos de los ensayos no permitió discriminar entre genotipos en ninguno de los índices aplicados. El ajuste de los índices de selección Smith, KN, ESIM, RESIM, y CPs no presentaron un patrón generalizado, por lo que el ajuste de estos índices debe ser priorizados para cada conjunto de datos. No se encontró una función que ponderara apropiadamente todas las 14 variables en estudio. Los ACPs, los análisis de varianza y las funciones lineales de los índices de desempeño indicaron que las variables con mayor peso en la variación fueron: **brixprom_1**, **pol_6**, **brixjugo_7**, **purity_8**, **peso5talks_2**, **diam_11**, **alt_10**, **fiber_14**, **flower_12**, **pith_13** y **oqueinter_4**. La variable **flower_12** fue de interés únicamente en la

diferenciación genética de variedades florecedoras. La selección de genotipos de caña de azúcar según las regresiones **Logit** fue función de los valores genéticos determinados únicamente por el TAH y BLUP. BLUP optimizó los procedimientos de selección de genotipos y puede ser aplicado de forma rutinaria en los procesos de evaluación y selección de genotipos en condiciones de cultivo en Guatemala. Por todo lo anterior, el procedimiento estadístico más apropiado para diferenciar genotipos de caña de azúcar fue el Mejor Estimador Lineal Insesgado (BLUP) a partir de la variable TAH. La utilización de esta metodología permitió la selección de 25 genotipos de caña de azúcar con un incremento promedio del 22% en el TAH (5.13 t) respecto a los testigos utilizados.



ABSTRACT

In sugarcane, with huge amount of populations under evaluation, it is necessary to observe all the important variables; therefore, the precise selection degree can be affected by the difficulty of an accurate simultaneous evaluation of all the involved variables in each genotype. Moreover, to reduce the dimensionality of all variables in one or few dimensions (indexes) would make possible to evaluate and classify the sugarcane genotypes in a more efficient fashion. Consequently, the present study was carried out with the next objectives: a) to evaluate nine methodologies for computing selection indices using phenotypic information in sugarcane in the Selection Stage III at the Guatemalan CENGICAÑA's breeding program; b) to determine the adjustment level of all indices models under study in order to establish its practical use; c) to identify the variables that better explains the performance variation of all sugarcane genotypes under study. To reach the objectives, four trials were settled-out in two different locations and two different seasons; seven data sets were collected-out, due that one of them was loosen by an non-scheduled burn in one trial. 513 sugarcane genotypes were contained by all the four trials. 14 phenotypic characterization variables were derived from 58 different measures. The information was analyzed by univariate descriptive statistics, "t" tests, conventional analysis of variance (ANOVA), linear mixed models analysis, linear regression analysis, multivariate analysis of variance (MANOVA), principal components analysis (PCA) and **Logit** regressions. Scores for nine selection indices were computed: Smith, Kempthorne y Nordskog, ESIM, RESIM, PC1, PC2, PC3, TAH and BLUP. It was determined that the TSH performance differences among genotypes were highly significant and showed that the genetic factor was determinant to the differentiation; nevertheless, the *genotype* \times *enviroment* showed the necessity of establishing comparative trials in diverse kind of environments. The poor biomass development in two of the trials did not allow to differentiate among genotypes in none of the all nine indices. The low adjustment of the selection indexes Smith, KN, ESIM, RESIM and PCs showed an erratic pattern; thus, the adjustment of these indexes, and probably other kind of indexes, must be prioritized according to every dataset. It was not possible to find a function that properly weighted all the 14 variables under study. Agreeing to the results in PCA, analysis of variance, linear regressions analysis and the linear functions of the selection indexes, **brixprom_1**, **pol_6**, **brixjugo_7**, **purity_8**, **peso5talks_2**, **diam_11**, **alt_10**, **fiber_14**, **flower_12**, **pith_13** and **oqueinter_4** were the

variables that most influenced the variation. The variable **flower_12** was useful in the genetic differentiation only among the flowering genotypes. In five of the seven dataset concurring to the **Logit** regressions, the selection of genotypes was a function only of TAH and BLUP. BLUP was able to optimize the selection procedures among sugarcane genotypes and can be routinely applied in the genotypes evaluation and in selection procedures. Consequently, the most proper statistical procedure to differentiate sugarcane genotypes was the Best Linear Unbiased Prediction (BLUP) derived from the TAH variable. Finally, it was determined that the application of this methodology allowed the selection of 30 of the 513 considered sugarcane genotypes with a genetic gain of 22% in TAH (5.13 t) as compared to the used check-varieties.



INDICE DE CONTENIDOS

Pág

1.	INTRODUCCIÓN	3
1.1	GENERALIDADES SOBRE LA CAÑA DE AZÚCAR	3
1.2	ASPECTOS GENERALES DE LA CITOGÉNICA DE LA CAÑA DE AZÚCAR	7
1.2.1	Citogenética molecular de la caña de azúcar	7
1.2.2	¿Cuál es el número básico de cromosomas de la caña de azúcar?	8
1.2.3	Los híbridos <i>S. officinarum</i> x <i>S. spontaneum</i>	8
1.3	EL PROGRAMA DE VARIEDADES DE CENGICAÑA	12
1.4	MEJORAMIENTO GENÉTICO DE LA CAÑA DE AZÚCAR.....	14
1.4.1	La interacción Genotipo × Ambiente ($G \times E$).....	16
1.4.2	Procedimientos de Selección	18
1.5	LOS ÍNDICES DE SELECCIÓN	21
1.5.1	Índice Función Discriminante Para la Selección de Plantas (Smith, 1936).....	23
1.5.2	Índices de selección por restricción de variables (<i>Restricted Selection Index</i>). (Kempthorne y Nordskog, 1959)	28
1.5.3	Las Componentes Principales (CPs) como índices de desempeño	31
1.5.4	Índice de selección basado en Eigenanalysis (<i>Eigenanalysis Selection Index Method</i> , ESIM) (Cerón-Rojas <i>et al.</i> , 2006)	34
1.5.4.1	La estimación de los <i>Eigenvalues</i> , <i>Eigenvectors</i> , ESIM y la Respuesta a la selección de ESIM (Cerón-Rojas <i>et al.</i> , 2006) ..	35
1.5.5	Índice de selección restrictivo por medio de “Eigenanalysis” (<i>Restricted Selection Index Method Based on Eigenanalysis</i> , RESIM) (Cerón-Rojas, <i>et al.</i> , 2008b).....	36
1.5.6	El mejor estimador lineal insesgado (Best linear unbiased predictor), BLUP	38
1.5.6.1	Modelos Lineales Mixtos, BLUE y BLUP	38
1.5.6.2	BLUP multi-carácter o MT-BLUP (<i>Multi Trait BLUP</i>)	41
1.5.7	Toneladas de Azúcar por Hectárea, TAH	45
1.6	PROBLEMA ESPECÍFICO DEL ESTADO III DEL PROGRAMA DE VARIEDADES DE CENGICAÑA	47
1.7	DISCUSIÓN Y JUSTIFICACIÓN DEL PROYECTO RESPECTO AL PROBLEMA ESPECÍFICO.....	47
2.	OBJETIVOS E HIPÓTESIS.....	53
2.1	OBJETIVOS	53
2.1.1	Objetivo General.....	54
2.1.2	Objetivos Específicos	54
2.2	HIPÓTESIS.....	54
3.	MATERIALES Y MÉTODOS.....	59
3.1	EXPERIMENTOS Y CONJUNTOS DE DATOS	59
3.2	MATERIAL VEGETAL	62
3.3	EVALUACIÓN Y SELECCIÓN.....	64
3.4	DISEÑO EXPERIMENTAL.....	66
3.5	VARIABLES DE ESTUDIO.....	67
3.6	ANÁLISIS DE LA INFORMACIÓN	68
3.6.1	Análisis general por variables	68
3.6.2	ANÁLISIS GENERAL DE VARIACIÓN COMBINADO	69
3.6.3	Heredabilidad	70
3.6.4	Análisis De Componentes Principales (ACP)	74
3.6.5	Análisis De Modelos Lineales Mixtos Para Los Scores De Las CPs	75

INDICE DE CONTENIDOS (Cont.)

Pág

3.6.6	Análisis Multidimensional De Variación (MANOVA)	76
3.6.7	Los Índices de Selección	77
3.6.7.1	Función Discriminante Propuesta Por Smith (Smith, 1936).....	77
3.6.7.2	Índice Restrictivo De Kempthorne Y Nordskog, KN (Kempthorne Y Nordskog, 1959)	78
3.6.7.3	<i>Eigenanalysis-Based Selection Index Method</i> (ESIM).....	79
3.6.7.4	<i>Restrictive Eigenanalysis-Based Selection Index Method</i> (RESIM)	79
3.6.7.5	CPs Como Índices De Desempeño.....	79
3.6.7.6	TAH, Índice t De Azúcar Por ha.....	80
3.6.7.7	BLUP (<i>Best Linear Unbiased Prediction</i>). El Mejor Estimador Lineal Insesgado	81
3.6.7.8	Base De Comparación Entre Índices: Competencia Del Fitomejorador	83
3.6.7.9	Regresión Logit.....	85
4	RESULTADOS	89
4.1	RESULTADOS GENERALES	89
4.2	BIOMASA.....	93
4.3	CONCENTRACIÓN DE SACAROSA	97
4.4	ANÁLISIS DE VARIANZA.....	100
4.5	FLORACIÓN	101
4.6	TAH (T DE AZÚCAR POR HA)	102
4.7	CORRELACIONES.....	103
4.8	ANÁLISIS GENERAL COMBINADO	107
4.9	ESTIMACIÓN DE LOS VALORES DE HEREDABILIDAD DE LAS CARACTERÍSTICAS EN ESTUDIO	108
4.10	ANÁLISIS DE COMPONENTES PRINCIPALES (ACP)	110
4.10.1	Medio Flor Plantía (medio_flor_plantía)	117
4.10.2	Medio No Flor Primera Soca (medio_no_flor_1ª_soca).....	121
4.10.3	Litoral No Flor Plantía (litoral_no_flor_plantía)	125
4.11	ANÁLISIS DE MODELOS LINEALES MIXTOS PARA LAS CPs PARA TODOS LOS DATOS	128
4.12	ANÁLISIS MULTIDIMENSIONAL DE VARIANZA (MANOVA)	135
4.13	ÍNDICES DE SELECCIÓN (DESEMPEÑO)	140
4.13.1	Índices de selección para el conjunto de datos medio flor plantía (medio_flor_plantía).....	140
4.13.1.1	Análisis de correlación entre índices	140
4.13.1.2	Índice "Smith" y "Restrictive Kempthorne and Nordskog Selection Index" (Índice KN)	141
4.13.1.3	<i>Eigenanalysis-based Selection Index Method</i> (ESIM) y <i>Restrictive Eigenanalysis-based Selection Index Method</i> (RESIM)	143
4.13.1.4	CPs como índices de desempeño para el conjunto de datos medio_flor_plantía	145
4.13.1.5	TAH (t De Azúcar Por ha) y BLUP (<i>Best Linear Unbiased Prediction</i>)	150
4.13.1.6	Regresión Logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos medio_flor_plantía	154
4.13.2	Índices de selección para el conjunto de datos medio no flor 1ª soca (medio_no_flor_1ª_soca)	158
4.13.2.1	Análisis de correlación entre índices	158
4.13.2.2	Índice "Smith" y "Restrictive Kempthorne and Nordskog Selection Index" (Índice KN)	159
4.13.2.3	<i>Eigenanalysis-based Selection Index Method</i> (ESIM) y <i>Restrictive Eigenanalysis-based Selection Index Method</i> (RESIM)	162
4.13.2.4	CPs como índices de desempeño para el conjunto de datos medio_no_flor_1ª_soca	164
4.13.2.5	TAH (t De Azúcar Por ha) y BLUP (<i>Best Linear Unbiased Prediction</i>)	168
4.13.2.6	Regresión logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos medio_no_flor_1ª_soca.....	172
4.13.3	Índices de selección para el conjunto de datos litoral no flor plantía (litoral_no_flor_plantía).....	174
4.13.3.1	Análisis de correlación entre índices	174
4.13.3.2	Índices "Smith" y "Restrictive Kempthorne And Nordskog Selection Index" (Índice KN)	175

INDICE DE CONTENIDOS (Cont.)

Pág

4.13.3.3	<i>Eigenanalysis-based Selection Index Method (ESIM) y Restrictive Eigenanalysis-based Selection Index Method (RESIM)</i>	178
4.13.3.4	CPs como Índices de Desempeño para el conjunto de datos litoral_no_flor_plantia	180
4.13.3.5	TAH (t De Azúcar Por ha) y BLUP (<i>Best Linear Unbiased Prediction</i>)	183
4.13.3.6	Regresión Logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos litoral_no_flor_plantia	186
4.13.4	Las regresiones Logit en términos generales	188
5	DISCUSIÓN	193
5.1	SOBRE LOS RESULTADOS GENERALES	193
5.1.1	Medidas estadísticas generales	193
5.1.2	Análisis de Varianza	196
5.2	CORRELACIONES	198
5.3	HEREDABILIDAD	201
5.4	ANÁLISIS DE COMPONENTES PRINCIPALES (ACP) GLOBAL	202
5.4.1	Consideraciones generales sobre el ACP para todos los conjuntos de datos	203
5.4.2	Análisis de modelos lineales mixtos para CPs para cada conjunto de datos	204
5.4.3	Análisis Combinado de Varianza para los CPs	205
5.5	MANOVA	206
5.6	ÍNDICES DE SELECCIÓN	206
5.6.1	Correlaciones entre Índices	207
5.6.2	Índices <i>Smith</i> y <i>Kempthorne</i> y <i>Nordskog</i>	207
5.6.3	ESIM y RESIM	210
5.6.4	CPs Como Índices de Desempeño	211
5.6.5	TAH y BLUP	212
5.6.6	Regresiones Logit	215
6	CONCLUSIONES	219
7	REFERENCIAS BIBLIOGRÁFICAS	227
ANEXO 1. LISTADO DE GENOTIPOS EN CADA EXPERIMENTO. LAS CELDAS SOMBREADAS EN CELESTE, CORRESPONDEN A LAS VARIEDADES TESTIGOS EN CADA EXPERIMENTO.		245
ANEXO 2. LISTADO DE LAS MEDICIONES REALIZADAS EN CADA UNO DE LOS CONJUNTOS DE DATOS.		258
ANEXO 3. ÍNDICES DE SELECCIÓN PARA TODOS LOS GENOTIPOS (ENTRY) EN SU RESPECTIVO CONJUNTO DE DATOS (TRIAL).		260
ANEXO 4. RESUMEN DEL ANÁLISIS DE VARIANZA PARA UN MODELO MIXTO, DONDE LA VARIABLE “VARIEDAD” (ENTRY) SE CONSIDERÓ FIJA, EL RESTO SE CONSIDERARON ALEATORIAS. VARIABLES DE RESPUESTA: SCORES DE 14 COMPONENTES PRINCIPALES (CPS).		266



Tabla 1.1. Proyección mundial sobre la producción de azúcar.	4
Tabla 1.2. Estadísticos descriptivos de la productividad de la agroindustria azucarera guatemalteca en las últimas siete zafras.	6
Tabla 1.3. Aspectos sobre variación cuantitativa para especies de reproducción sexual y vegetativa.	17
Tabla 1.4. Algunas características que debe reunir una variedad de caña de azúcar.	48
Tabla 3.1. Condiciones climáticas, valores promedio mensuales, en las áreas de experimentación durante dos ciclos de evaluación, en el período 01/01/2014 Al 31/03/2016.	62
Tabla 3.2. Número de genotipos en evaluación por experimento.	63
Tabla 3.3. Listado de requisitos del cliente para las variedades que libera CENGICAÑA En Guatemala.	66
Tabla 3.4. Variables evaluadas, correspondencia y tipo de variable en R.	67
Tabla 3.5. Listado de variables para las cuales se estimó la heredabilidad y el cálculo de los índices de selección.	74
Tabla 4.1. Medidas generales descriptivas para las variables en estudio.	90
Tabla 4.2. Medidas descriptivas de 15 variables por conjunto de datos. Para cada genotipo estudiado se utilizaron muestras de diez tallos.	91
Tabla 4.3. Pruebas de <i>t</i> para dos poblaciones independientes y para quince variables, agrupadas según patrón de floración (flor y no_flor) y estrato de evaluación (litoral y medio).	95
Tabla 4.4. Resumen de los análisis de variación para las variables en estudio.	101
Tabla 4.5. Valores de correlación <i>R</i> de Pearson y nivel de significación	103
Tabla 4.6. Resumen de los resultados del Análisis de Varianza (ANOVA) combinado para la variable <i>t</i> de azúcar por ha (TAH).	107
Tabla 4.7. Valores estimados de heredabilidad para cada variable en estudio en tres experimentos con evaluaciones realizadas durante dos años (litoral_flor, medio_flor, medio_no_flor) y un año (litoral_no_flor).	109
Tabla 4.8. <i>Eigenvalues</i> (varianzas) y su participación porcentual en cada CP para todos los datos. .	111
Tabla 4.9. Cargas (<i>eigenvectors</i> o <i>loadings</i>) de cada variable original con respecto a las primeras seis CPs.	112
Tabla 4.10. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el ensayo medio_flor_plantía.	117
Tabla 4.11. Cargas (<i>Eigenvectors</i> o <i>loadings</i>) de cada variable original para las primeras cuatro CPs del conjunto de datos medio_flor_plantía.	118
Tabla 4.12. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el ensayo medio_no_flor_1ª_soca.	121
Tabla 4.13. Cargas (<i>Eigenvectors</i> o <i>loadings</i>) de cada variable original con respecto a las cuatro primeras CPs para el conjunto de datos medio_no_flor_1ª_soca.	122
Tabla 4.14. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el conjunto de datos litoral_no_flor_plantía.	125

INDICE DE TABLAS (Cont.)

Pág

Tabla 4.15. Cargas (<i>Eigenvectors</i> o <i>loadings</i>) de cada variable original con respecto a las cuatro primeras CPs para el conjunto de datos litoral_no_flor_plantía.	126
Tabla 4.16. Análisis de Varianza combinado para las variables CP1, CP2, CP3 y CP4.	128
Tabla 4.17. Resumen de Análisis de varianza Tipo III para 14 CPs en tres conjuntos de datos para el factor “variedad” (ENTRY). Aproximación Satterthwaite para los grados de libertad.....	129
Tabla 4.18. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos medio_flor_plantía.....	131
Tabla 4.19. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos medio_no_flor_1a_soca.....	132
Tabla 4.20. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos litoral_no_flor_plantía.	134
Tabla 4.21. Resumen de los Análisis Multidimensionales de Varianza (MANOVA). Solo para la variable de clase (independiente) VARIEDAD (ENTRY).	139
Tabla 4.22. Valores de correlación r de Pearson y probabilidad de F para probar la hipótesis $\beta_1 = 0$ para nueve índices de desempeño en el conjunto de datos medio_flor_plantía.....	141
Tabla 4.23. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_flor_plantía, ordenados descendientemente según los valores del índice de Smith.	142
Tabla 4.24. Resumen de las Salidas de RINDSEL para los índices Smith y KN y la jerarquización de la fracción superior (10%).	143
Tabla 4.25. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_flor_plantía.....	144
Tabla 4.26. Resumen de las Salidas de RINDSEL (Pérez-Elizalde <i>et al.</i> , 2014) para los índices ESIM y RESIM para el conjunto de datos medio_flor_plantía.	145
Tabla 4.27. Resumen de indicadores para los índices CP1, CP2, y CP3.	146
Tabla 4.28. Ordenación de los genotipos (fracción del 25%) por valores de las CPs 1, 2 y 3 (CP1, CP2, y CP3).....	147
Tabla 4.29. Fracción superior del 25% (8 variedades) inicialmente seleccionada por la CP1, seguida de la selección del 10% de la fracción superior, obtenidas por las CP2 y CP3.	149
Tabla 4.30. Resumen de indicadores del índice TAH (t de azúcar por ha).....	150
Tabla 4.31. Ordenamiento descendente de la fracción superior (23%) de genotipos según los valores de BLUP.....	154
Tabla 4.32. Valores de Pseudo r^2 de McFadden de una regresión Logit para el modelo general $\ln[p/(1-p)] = a + bX + e$ para el conjunto de datos medio_flor_plantía.....	155
Tabla 4.33. Correlación de Pearson y nivel de significación entre los valores de los nueve índices obtenidos para el conjunto de datos “medio_no_flor_1ª_soca”.	159
Tabla 4.34. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_no_flor_1ª_soca.....	160
Tabla 4.35. Resumen de las salidas de RINDSEL para los índices “Smith” y “KN” para el ensayo medio_no_flor_1ª_soca.	161
Tabla 4.36. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_no_flor_1ª_soca.....	163

INDICE DE TABLAS (Cont.)

Pág

Tabla 4.37. Resumen de las Salidas de RINDSEL para los índices ESIM y RESIM para el conjunto de datos “medio_no_flor_1ª_soca”.....	164
Tabla 4.38. Resumen de indicadores para los índices CP1, CP2 y CP3 para el conjunto de datos medio_no_flor_1ª_soca.....	165
Tabla 4.39. Clasificación de los genotipos (fracción del ~25%) según los <i>scores</i> de CP1, CP2 y CP3.	166
Tabla 4.40. Fracción superior de ~25% inicialmente seleccionada por la CP1 para el conjunto de datos medio_no_flor_1ª_soca. Separación de medias mediante el método de Tukey para la CP1. En verde, genotipos resultantes de la selección anidada a partir de las CP1, CP2 y CP3.	168
Tabla 4.41. Resumen de indicadores de los índices TAH para el conjunto de datos medio_no_flor_1ª_soca.....	168
Tabla 4.42. Frecuencias de genotipos evaluados en los tres conjuntos de datos según sus valores de TAH.	169
Tabla 4.43. Ordenamiento descendente de la fracción superior (~25%) de genotipos según los valores de BLUP para medio_no_flor_1ª_soca.	171
Tabla 4.44. Frecuencias de distribución de BLUP por cuartiles.....	171
Tabla 4.45. Valores de Pseudo r^2 de McFadden de una regresión Logit para el modelo general $\ln[p/(1-p)] = a + bX + e$ para el conjunto de datos medio_no_flor_1ª_soca.....	172
Tabla 4.46. Correlación de Pearson y Significación estadística entre los valores de los índices obtenidos para el conjunto de datos litoral_no_flor_plantía.....	175
Tabla 4.47. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo litoral_no_flor_plantía.....	176
Tabla 4.48. Resumen de las Salidas de RINDSEL para los índices Smith y KN para el conjunto de datos litoral_no_flor_plantía.	178
Tabla 4.49. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo litoral_no_flor_plantía.....	179
Tabla 4.50. Resumen de las Salidas de RINDSEL para los índices ESIM y RESIM para el conjunto de datos medio_no_flor_1ª_soca.....	180
Tabla 4.51. Resumen de indicadores para los índices CP1, CP2 y CP3 para el conjunto de datos litoral_no_flor_plantía.	181
Tabla 4.52. Resumen de <i>scores</i> para los índices CP 1, 2 y 3 para el conjunto de datos litoral_no_flor_plantía del ~25% de la fracción superior.....	182
Tabla 4.53. Fracción superior del 25% (10 variedades) seleccionada por la CP1. Comparación por pares ajustada por el método de Tukey y ordenación por las CP2 y CP3 para la selección final (~10% de la fracción superior, en verde).	183
Tabla 4.54. Resumen de indicadores de los índices TAH para el conjunto de datos litoral_no_flor_plantía	184
Tabla 4.55. Ordenamiento descendente de la fracción superior (~25%) de genotipos según los valores de TAH y BLUP. Separación de medias ajustada por la metodología de Tukey.....	186



INDICE DE FIGURAS

Pág

Figura 1.1. Producción, consumo y existencias proyectadas de azúcar en t (OECD/FAO, 2016).	5
Figura 1.2. El cruzamiento entre <i>S. officinarum</i> y <i>S. spontaneum</i>	9
Figura 1.3. Pedigrí de POJ 2878 y POJ 2725 (Purseglove, 1972; Sreenivasaan <i>et al.</i> , 1987).....	9
Figura 1.4. Cultivar moderno de caña de azúcar, CG98-46 (Híbrido <i>Saccharum</i> spp.).....	10
Figura 1.5. Esquema de trabajo para el mejoramiento de caña de azúcar en el Programa de Variedades de CENGICAÑA.	14
Figura 1.6. Frecuencias de distribución según la función discriminante de Fisher (Smith, 1936).....	25
Figura 3.1. Esquema de ordenación de los conjuntos de datos sujetos a evaluación.....	60
Figura 3.2. Ubicación geográfica de los cuatro experimentos en evaluación: a) estrato medio_flor. B) estrato litoral_flor. C) estrato medio_no_flor. D) estrato litoral_no_flor (Google Earth, SF).	61
Figura 4.1. a) Número de genotipos (frecuencias) según la variable peso5stalk_2. b) Número de genotipos según el contenido aparente de sacarosa (pol_6).	94
Figura 4.2 Número de genotipos para la variable peso de 5 tallos (peso5stalk_2).....	96
Figura 4.3. Contenido de fibra (%) (fiber_14), por conjunto de datos.	97
Figura 4.4. Promedio de valores del pol % caña (pol_6) por conjunto de datos.....	98
Figura 4.5. Promedios °brix en % medidos en el campo (brixprom_1).	98
Figura 4.6. Número de genotipos según °brix medidos en el laboratorio de jugos (brixjugo_7).	99
Figura 4.7. Promedios para la variable (índice) TAH para cada uno de los experimentos evaluados.	102
Figura 4.8. Regresión entre variables relacionadas con la concentración de azúcar: a) Pol % Caña Vs °Brix Campo (pol_6 vs brixprom_1) ($r^2 = 0.4626$, ***, total de datos) y b) °Brix Jugo Vs °Brix en campo (brixjugo_7 vs brixprom_1) ($r^2 = 0.4456$, ***, total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza	104
Figura 4.9. Regresión entre variables relacionadas con la concentración de azúcar: a) Pol % Caña vs °Brix campo ($r^2 = 0.476$, *** total de datos) y b) °Brix Jugo en laboratorio Vs °brix campo ($r^2 = 0.504$, ***, total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza.	105
Figura 4.10. Regresión entre las variables Peso de cinco tallos vs. Diámetro según: a) el estrato de evaluación (litoral y medio) ($r^2 = 0.24$, ***, total de datos); y b) según el tipo de Floración (flor y no_flor) ($r^2 = 0.34$, ***, total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza	106
Figura 4.11. Regresión Pol % caña vs población de tallos por metro lineal ($r^2 = 0.202$, ***).....	106
Figura 4.12. Varianzas por CP al considerar todas las observaciones de los siete conjuntos de datos.	111
Figura 4.13. Distribución de los siete conjuntos de datos del estudio según las dos primeras CPs (Dim1 vs Dim2) que acumularon el 40.6% de la variabilidad.....	113
Figura 4.14. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c).....	114
Figura 4.15. Distribución de todos los siete conjuntos de datos según tres CPs (CP1 y CP2 y CP3) para siete conjuntos de datos.	115
Figura 4.16. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c).....	119
Figura 4.17. a) Gráfico De Dispersión (Biplot) CP1 y CP2; b) Gráfico De Dispersión (Triplot) CP1, CP2 y CP3 Para El Conjunto De Datos medio_flor_plantía.	120

INDICE DE FIGURAS (Cont.)

Pág

Figura 4.18. a) Gráfico de dispersión Biplot para las CP1 y CP2, y b) Gráfico de dispersión Triplot para las CP1, CP2 y CP3. Conjunto de datos medio_no_flor_1ª_soca.	123
Figura 4.19. Contribución en porcentaje de las 14 variables a las CP1 (a), CP2 (b) y CP3(c) para el conjunto de datos medio_no_flor_1ª_soca.	124
Figura 4.20. a) Gráfico de dispersión Biplot para las CP1 y CP2, y b) Gráfico de dispersión Triplot para las CP1, CP2 y CP3. Conjunto de datos litoral_no_flor_plantía.	126
Figura 4.21. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c), para el conjunto de datos litoral_no_flor_plantía.	127
Figura 4.22. Gráfico “cuantil-cuantil” (QQ-plot) para tres conjuntos de datos: a) litoral_no_flor_plantía. b) medio_flor_plantía. c) medio_no_flor_1ª_soca para la variable (índice) CP1.	129
Figura 4.23. Comparación por pares ajustada por el método de Tukey para el conjunto de datos medio_flor_plantía para la CP1.	130
Figura 4.24. Comparación por pares ajustada por el método de Tukey para el conjunto de datos medio_no_flor_1ª_soca para la CP1.	132
Figura 4.25. Comparación por pares ajustada por el método de Tukey para el conjunto de datos litoral_no_flor_plantía para la CP1.	134
Figura 4.26. Matriz de diagramas de dispersión y “r” de Pearson entre cuatro variables relacionadas con la concentración de sacarosa para el total de datos (Los niveles de significación estadística de estas regresiones se presentaron en la Tabla 4.5.).	136
Figura 4.27. Matriz de diagramas de dispersión y “r” de Pearson entre cuatro variables relacionadas con acumulación de biomasa para el total de datos (los niveles de significación estadística se presentaron en la Tabla 4.5).	137
Figura 4.28. Matriz de diagramas de dispersión y “r” de Pearson entre cinco variables relacionadas con características de calidad del tallo para el total de datos (significación estadística en la Tabla 4.5).	138
Figura 4.29. Gráficos de densidades (a) y cuantiles (b) para la CP1 del conjunto de datos medio_flor_plantía.	148
Figura 4.30. Gráficos de residuos (a) y de cuantiles (b) sobre la distribución de los datos de las observaciones para la variable/índice TAH (t de azúcar por ha) en el conjunto de datos “litoral_flor_1ª_soca”.	151
Figura 4.31. Distribución de las frecuencias de los BLUPs.	152
Figura 4.32. Relación BLUP con las medias de TAH.	153
Figura 4.33. Gráficos de Regresión lineal Logit para la relación entre la base de comparación de selección por competencia (“Si” o “No”) con seis índices de selección: a) TAH (t de azúcar por ha); b) BLUP de TAH; c) CP2; d) ESIM; e) CP3; y f) Smith.	157
Figura 4.34. Propiedades sobre normalidad de la serie de datos medio_no_flor_1ª_soca.	167
Figura 4.35. a) Distribución de frecuencias absolutas de los BLUPs. b) BLUPs ajustados a la media de TAH.	170
Figura 4.36. Gráficos de las funciones logit en el conjunto de datos medio_no_flor_1ª_soca para los índices: a) TAH , b) BLUP, c) ESIM, d) CP1, e) KN y f) Smith.	173

INDICE DE FIGURAS (Cont.)

Pág

Figura 4.37. Gráficos de densidades (a) y cuantiles (b) para la variable CP1 del conjunto de datos litoral_no_flor_plantía.	182
Figura 4.38. Ajuste de BLUP con respecto de la media de TAH (a) y distribución de las frecuencias de BLUP (b).	185
Figura 4.39. Gráficos de las funciones Logit para los índices: a) BLUP, b) TAH, c) RESIM, d) SMITH, e) PC1, Y f) PC2.	187
Figura 4.40. Gráficos de Regresión lineal Logit para la relación entre la base de comparación de selección por competencia (“Sí” o “No”) con los índices TAH y BLUP según el estrato en estudio (a y b) o el patrón de floración (c y d) y ESIM según estrato de evaluación (e) o el patrón de floración (f).	189







1. INTRODUCCIÓN



1. INTRODUCCIÓN

1.1 GENERALIDADES SOBRE LA CAÑA DE AZÚCAR

La caña de azúcar es una importante fuente de alimento y bioenergía. Es relevante en las economías de muchos países en el trópico y el sub-trópico. En Guatemala, el azúcar proveniente de la caña de azúcar representa el 9% de las exportaciones totales del país (Banco de Guatemala, 2015). Los atributos que otorgan a la caña de azúcar un valor económico importante son: su alta productividad; uso eficiente de los insumos agrícolas; y, puede ser procesada localmente para la adición de valor, mediante la producción de azúcar, melazas, alcohol y energía, todos susceptibles de almacenamiento y transporte (Moore *et al.*, 2014).

La caña de azúcar es un pasto grande y perenne que evolucionó en condiciones de alta luminosidad, altas temperaturas y grandes cantidades de agua, por lo que se adapta a las regiones ubicadas al norte y al sur de los 35° desde el ecuador (Moore *et al.*, 2014). La superficie estimada de producción de la caña de azúcar es de más de 24 millones de ha (FAO, 2015) y la estimación de la producción mundial para el período 2016/2017 es de poco más de 170 millones de t de azúcar (FAS/USDA, 2017). La caña de azúcar es el cultivo de mayor producción de biomasa (caña) en el mundo, comparado con el resto de cultivos de importancia económica (FAO, 2015), esto sin considerar el contenido de materia seca. Por otro lado, la caña de azúcar provee un tercio de la energía proveniente de plantas consumida por los humanos (152 Kcal/cápita/día), solamente por detrás del arroz y el trigo (Moore *et al.*, 2014).

A escala mundial la caña de azúcar tiene una importancia relevante, relacionada con su alta capacidad de producción de biomasa. El azúcar en el mundo se obtiene principalmente de dos cultivos fuentes: la caña de azúcar y la remolacha azucarera. Esta última se cultiva

principalmente en países desarrollados, mayormente en Europa y Estados Unidos (Quemé y Melgar, 2016). Por otra parte, la caña de azúcar se cultiva en muchos otros países, principalmente del trópico y del subtrópico (Estados Unidos, México, Centroamérica, Colombia, Ecuador, Venezuela, Argentina, Brasil, Sudáfrica, Australia, Tailandia, India, etc.). La relación de producción de azúcar proveniente de la caña de azúcar versus la remolacha azucarera es del 87.5% al 12.5% (Tabla 1.1). El azúcar es una relevante fuente de energía (*commodity*) y, como tal, se comercializa en el mundo (Moore *et al.*, 2014). La proyección mundial para el 2020 indica que la mayor proporción de producción de azúcar será del 87% proveniente de caña de azúcar y el resto de remolacha azucarera (OECD/FAO, 2016).

Tabla 1.1. Proyección mundial sobre la producción de azúcar.

	Dimensional	Promedio 2013-2015	2016	2017	2018	2019	2020
Remolacha azucarera							
Producción	t (millones)	257.0	268.2	271.9	274.2	276.8	278.9
Área	ha (millones)	4.4	4.5	4.5	4.5	4.6	4.6
Rendimiento	t/ha	58.63	58.99	59.79	60.31	60.71	61.16
Uso para combustible	t (millones)	12.8	12.7	10.0	10.5	10.4	10.4
Caña de Azúcar							
Producción	t (millones)	1811.6	1850.3	1875.1	1908.5	1943.3	1976.0
Área	ha (millones)	26.9	27.2	27.3	27.7	28.0	28.3
Rendimiento	t/ha	67.37	68.08	68.57	68.98	69.41	69.74
Uso para combustible	t (millones)	374.3	409.5	414.0	420.0	430.3	440.1
Azúcar en General							
Producción	t (millones) tq	170.8	173.7	177.8	181.9	187.2	191.4
Consumo	t (millones) tq	166.8	172.9	175.3	178.2	181.8	185.7
Precio de azúcar cruda ⁽¹⁾	USD/t	323.0	325.3	330.4	337.6	335.8	330.9
Precio de azúcar blanca ⁽²⁾	USD/t	408.0	398.6	388.3	402.5	405.8	407.0
Precio, HFCS ⁽³⁾	USD/t	539.7	472.1	450.8	457.6	470.1	477.2

(1) Precio mundial de azúcar cruda para 2016 (ICE contract No11 nearby, October/September)

(2) Precio de azúcar refinada para 2016 (White Sugar Futures Contract No. 407, Euronext market, Liffe, London, Europe, October/September)

(3) Lista de precios de ventas completas en los Estados Unidos para 2016 (HFCS-55, -high fructose corn syrup- October/September)
tq= *tel quel* (sobre la base de producciones netas). Fuente: OECD/FAO, 2016.

Por otro lado, si bien los precios son fluctuantes, se espera un crecimiento en la producción de cerca de un 2.1% anual, y un crecimiento sostenido de la demanda en cerca de un 2% anual (OECD/FAO, 2016). En todo caso, se asume que en el nivel mundial la oferta será mayor que la demanda (Figura 1.1).

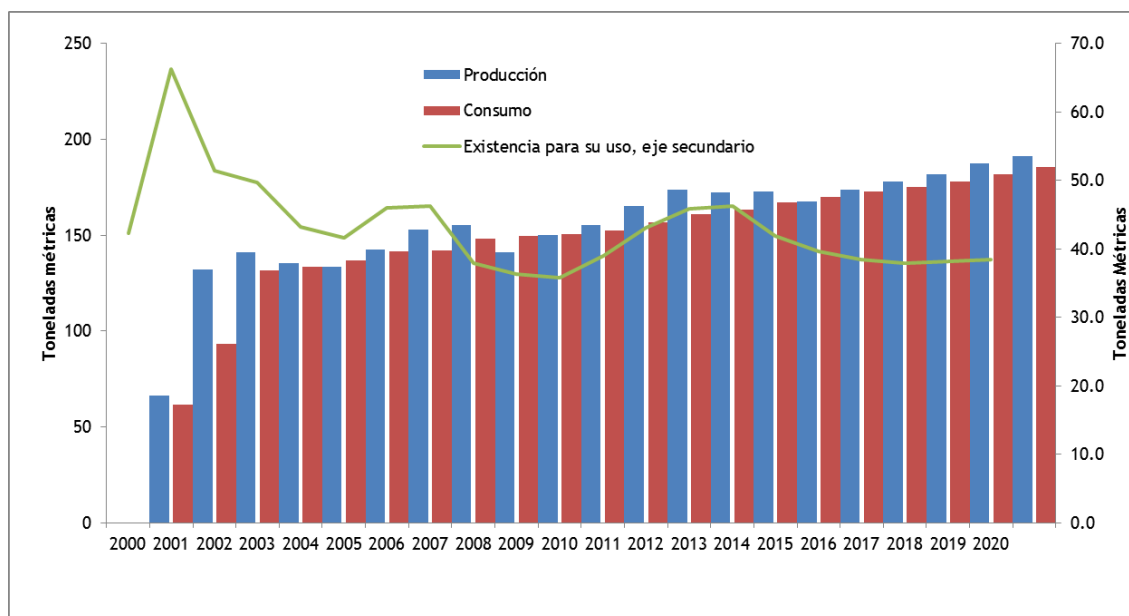


Figura 1.1. Producción, consumo y existencias proyectadas de azúcar en t (OECD/FAO, 2016).

En Guatemala, el azúcar se produce exclusivamente a partir de caña de azúcar. En la zafra de 2016-2017 se cosecharon cerca de 250,000 ha, de las cuales se obtuvieron 184,728 t (CENGICAÑA, 2011-2017). La productividad promedio de caña de la agroindustria para la zafra 2016-2017 fue de 101.96 t ha^{-1} , el *Pol % Caña* medio (medida de la concentración del azúcar en la caña) fue de 13.10%, y la combinación de ambas variables, en promedio, propició la productividad de 10.86 t de azúcar por ha (CENGICAÑA, 2011-2017). Con estos valores de productividad, y considerando la productividad de otros años, Guatemala ocupa el tercer lugar en productividad de azúcar en el mundo, solamente por detrás de Colombia y Suazilandia (Melgar, 2012).

La evolución de la productividad de la agroindustria azucarera guatemalteca en las últimas siete zafas (Tabla 1.2) muestra cómo el rendimiento de caña se ha aumentado de 90 t ha^{-1} para la zafra 2010-2011 a 102 t ha^{-1} para la zafra 2016-2017. Por otro lado, la producción de azúcar por t de caña se ha mantenido más o menos constante ($\sim 10.5 \text{ kg t}^{-1} \text{ caña}$), en parte debido a que los cultivares de caña y la tecnología de extracción tienden a concentrar más o menos la misma cantidad de azúcar en los culmos. Esto indica que este carácter en particular es relativamente difícil de incrementar, lo que obliga a establecer un equilibrio entre la producción de biomasa, de la concentración de azúcar y de la distancia a la fábrica para el transporte.

Tabla 1.2. Estadísticos descriptivos de la productividad de la agroindustria azucarera guatemalteca en las últimas siete zafra.

Zafra	Variable	Media	D.E.	Mín	Máx	Mediana	Q1	Q3
Zafra 2010-2011	Caña por hectárea (t)	90.5	25.8	10.7	340.3	90.1	73.3	106.6
	azúcar por tonelada caña (kg t ⁻¹)	10.9	1.1	5.2	14.8	10.8	10.2	11.5
	TAH	9.8	2.9	1.3	42.0	9.7	7.9	11.5
	°Brix	ND	ND	ND	ND	ND	ND	ND
	Pol % Caña	13.3	1.6	1.1	16.9	13.5	12.7	14.1
	Fibra %	ND	ND	ND	ND	ND	ND	ND
Zafra 2011-2012	Caña por hectárea (t)	101.1	37.1	1.3	2682.3	100.4	83.5	117.6
	azúcar por tonelada caña (kg t ⁻¹)	10.6	1.0	5.4	17.3	10.7	10.0	11.3
	TAH	10.7	4.0	0.2	277.1	10.6	8.7	12.5
	°Brix	ND	ND	ND	ND	ND	ND	ND
	Pol % Caña	13.1	1.7	1.0	17.0	13.2	12.6	13.9
	Fibra %	ND	ND	ND	ND	ND	ND	ND
Zafra 2012-2013	Caña por hectárea (t)	104.4	27.6	10.3	473.4	103.3	86.9	120.1
	azúcar por tonelada caña (kg t ⁻¹)	10.5	1.2	2.1	25.0	10.6	9.9	11.2
	TAH	10.9	3.0	1.0	54.8	10.8	9.0	12.6
	°Brix	ND	ND	ND	ND	ND	ND	ND
	Pol % Caña	12.9	2.3	1.0	28.0	13.3	12.6	14.0
	Fibra %	ND	ND	ND	ND	ND	ND	ND
Zafra 2013-2014	Caña por hectárea (t)	104.0	28.9	1.5	690.1	103.3	86.2	121.1
	azúcar por tonelada caña (kg t ⁻¹)	10.3	1.3	2.8	24.2	10.3	9.6	11.0
	TAH	10.6	3.0	0.2	76.8	10.5	8.7	12.3
	°Brix	ND	ND	ND	ND	ND	ND	ND
	Pol % Caña	12.5	2.7	1.0	29.8	12.9	12.4	13.6
	Fibra %	ND	ND	ND	ND	ND	ND	ND
Zafra 2014-2015	Caña por hectárea (t)	105.9	27.9	0.1	472.4	105.0	88.0	122.6
	azúcar por tonelada caña (kg t ⁻¹)	10.7	1.1	2.6	16.4	10.7	10.0	11.4
	TAH	11.3	3.1	0.0	51.9	11.1	9.3	13.1
	°Brix	ND	ND	ND	ND	ND	ND	ND
	Pol % Caña	13.1	2.0	1.0	18.6	13.3	12.7	14.0
	Fibra %	ND	ND	ND	ND	ND	ND	ND
Zafra 2015-2016	Caña por hectárea (t)	108.4	33.2	8.9	1469.9	106.7	88.1	126.5
	azúcar por tonelada caña (kg t ⁻¹)	10.2	1.2	4.1	14.6	10.3	9.6	11.0
	TAH	11.0	3.4	1.1	167.1	10.8	8.9	12.9
	°Brix	16.4	5.1	1.7	229.0	17.9	15.9	18.9
	Pol % Caña	12.6	2.4	1.0	18.3	12.9	12.3	13.6
	Fibra %	13.9	3.1	1.3	138.5	13.7	13.0	14.6
Zafra 2016-2017	Caña por hectárea (t)	102.0	28.1	8.8	591.8	101.0	84.6	118.6
	azúcar por tonelada caña (kg t ⁻¹)	10.7	1.0	1.1	15.8	10.7	10.1	11.4
	TAH	10.9	3.0	1.0	58.5	10.8	9.0	12.6
	°Brix	16.4	4.9	1.9	24.6	18.2	16.0	19.1
	Pol % Caña	13.0	1.7	1.1	17.2	13.2	12.5	13.8
	Fibra %	13.5	1.4	1.1	18.9	13.6	12.9	14.3

ND = No disponible; TAH = Toneladas de Azúcar por Hectárea; Pol % Caña = Concentración Aparente de Sacarosa. CENGICANA, 2011-2017.

Wagner (2007) indica que la caña de azúcar fue introducida a Guatemala, y se inició su cultivo hacia 1536. La misma autora señala que los primeros trapiches fueron construidos en el valle de Guatemala y el municipio de San Jerónimo (Baja Verapaz). Estos trapiches estuvieron operados, originalmente, por órdenes religiosas católicas. La fundación de ingenios

(fábricas) de azúcar se iniciaron hacia mediados del siglo XX, época en la que se inició también la exportación de azúcar en Guatemala (Wagner, 2007).

Actualmente, el azúcar en Guatemala es producida por diez ingenios azucareros, los cuales están aglutinados en la Asociación de Azucareros de Guatemala (AZASGUA), organización fundada en 1957. El Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar, CENGICANÁ, es una organización perteneciente a AZASGUA.

1.2 ASPECTOS GENERALES DE LA CITOGÉNÉTICA DE LA CAÑA DE AZÚCAR

Con respecto a los orígenes de la caña de azúcar, se sabe que tuvieron lugar en el sur de Asia durante el período cretáceo. Se cree que un género ancestral de la caña de azúcar se originó a partir de especies de los géneros *Erianthus* y *Mischantus* (James, 2004). Taxonómicamente la caña de azúcar se ubica en la tribu *Andropogonae*, la familia *Poaceae* y es parte del género *Saccharum*. El “Taxonomic Information System” (ITIS, 2017) lista un total de 14 especies aceptadas, con dos variedades de *S. brevibarbe*. Actualmente la literatura acepta seis especies (Ming *et al.*, 2006). Solamente *S. spontaneum* y *S. robustum* son consideradas especies silvestres, las otras cuatro son especies cultivadas (*S. barberi*, *S. edule*, *S. officinarum* y *S. sinense*) (Ming *et al.*, 2006). Debido a que ninguna de las cuatro especies cultivadas sobrevive en la naturaleza por sí sola, se las han considerado como *cultivares*. Así mismo, *S. barberi*, *S. sinense* y *S. edule* se cree que son de origen interespecífico o intergenérico (D’Hont *et al.*, 1998). Por otro lado, entre las seis especies indicadas, se han producido múltiples hibridaciones (artificiales o naturales), lo que ha derivado en una complicación mayor en términos de la ubicación taxonómica de las mismas. Esta amplia hibridación ha llevado a la generación de mezclas de cañas de azúcar euploides y aneuploides (Ming *et al.*, 2006).

1.2.1 Citogenética molecular de la caña de azúcar

Estudios de citogenética molecular (D’Hont *et al.*, 1998; Grivet *et al.*, 2004; Edmé *et al.*, 2005; Babu *et al.*, 2006; Piperidis *et al.*, 2010) y de mapeo genético (Da Silva *et al.*, 1993; Al-Janabi *et al.*, 1993; Grivet *et al.*, 1994) han establecido el tamaño aproximado del genoma de *S. spontaneum* entre 3.05 y 5.31 pg (picogramos, 1 pg=987 Mbp), en tanto que, el tamaño del genoma para *S. officinarum* oscila entre 6.32 y 6.66 pg. Algunos cultivares comerciales

(híbridos *Saccharum* spp) de Canal Point tienen un genoma que oscila entre 6.30 y 7.5 pg (Edmé *et al.*, 2005). Los cultivares modernos de caña de azúcar exhiben alrededor del 70 al 80% de los cromosomas enteramente derivados de *S. officinarum*; y entre el 10 y el 20% proveniente de *S. spontaneum*, y muy pocos cromosomas producto de la recombinación interespecífica (Ming *et al.*, 2006; Le Cunff *et al.*, 2008).

1.2.2 ¿CUÁL ES EL NÚMERO BÁSICO DE CROMOSOMAS DE LA CAÑA DE AZÚCAR?

Sreenivasan *et al.* (1987) hallaron que las propuestas del número básico son: $X = 5, 6, 8, 10$, y 12 . En *S. officinarum*, el total de cromosomas es $2n = 10x = 80$. Aquellos clones con un mayor número de cromosomas son considerados atípicos o híbridos. Para *S. officinarum*, con citotipos principales de $2n = 60-80$, el número básico más probable es $x = 10$ (D'Hont *et al.*, 1998; Butterfield y D'Hont, 2001; Ming *et al.*, 2006).

S. spontaneum exhibe un amplio rango en el número de cromosomas, de $2n = 36$ a $2n = 128$; con cinco citotipos mayores $2n = 64, 80, 96, 112$ y 128 (Ming *et al.*, 2006). D'Hont *et al.* (1998) establecieron mediante fluorescencia, usando los genes 18s-25s rDNA y 5S rDNA, la ubicación física de estos en los cromosomas de diferentes citotipos de *S. spontaneum*, determinando que el número total de cromosomas fue proporcional a los números de sitios de rDNA mapeados; así, se derivó que el número básico de juegos de cromosomas para *S. spontaneum* es de $x = 8$.

1.2.3 Los híbridos *S. officinarum* x *S. spontaneum*

La hibridación interespecífica en el género *Saccharum* (Figura 1.2) fue iniciada por mejoradores holandeses hacia 1885. Producto de este trabajo, se obtuvieron los cultivares POJ2725 y POJ2878, los cuales han contribuido significativamente como progenitores para los cultivares modernos de caña de azúcar alrededor del mundo (Purseglove, 1972). Los híbridos interespecíficos que involucran *S. officinarum* como el parental materno y *S. spontaneum* como parental paterno tienen un esquema triple (AAB) de cromosomas (Figura 1.3); por ejemplo, cruzando *S. officinarum* ($2n = 10x = 80$) por *S. spontaneum* ($2n = 8x = 112$) puede producir híbridos con $2n = 136$ cromosomas ($40 + 40$ de *S. officinarum*, más 56 de *S. spontaneum*, $2n + n$) (Sreenivasan *et al.*, 1987). Estos híbridos son caracterizados por su bajo contenido de azúcar, tallos delgados, alto contenido de fibra, gran capacidad de rebrote, y

elevados niveles de resistencia a factores bióticos y abióticos. Para reducir los efectos negativos de *S. spontaneum*, y retener la alta concentración de sacarosa de *S. officinarum*, se realizaron varios retrocruzamientos entre los híbridos interespecíficos y clones de *S. officinarum* (Figura 1.3). Esto condujo a la *nobilización* de lo que hoy denominamos híbridos *Saccharum* spp. (Sreenivasan *et al.*, 1987).



Saccharum officinarum



Saccharum spontaneum

Figura 1.2. El cruzamiento entre *S. officinarum* y *S. spontaneum*.

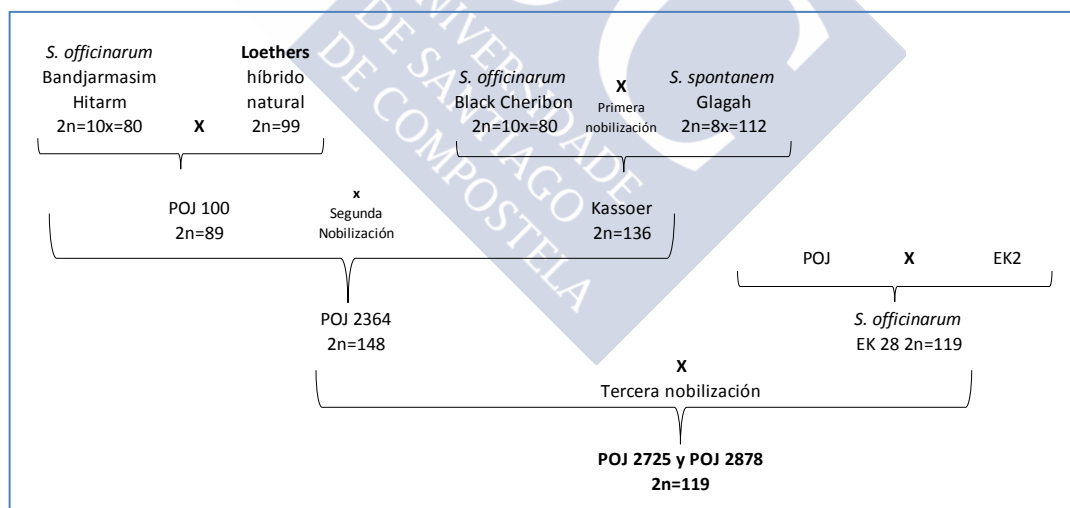


Figura 1.3. Pedigrí de POJ 2878 y POJ 2725 (Purseglove, 1972; Sreenivasaan *et al.*, 1987).

Al realizar el retrocruzamiento, el resultado fue, una vez más, una progenie dotada con $2n + n$ gametos. Generaciones subsecuentes de retrocruzas presentaron reducción de gametos (Figura 1.3). Este proceso condujo a la pérdida de cromosomas en las generaciones resultantes: la aneuploidía (Sreenivasaan *et al.*, 1987; Butterfield y D'Hont, 2001; 'Hont,

1998). Por esta razón, dependiendo de la cruce y del híbrido, los cultivares modernos de caña de azúcar son altamente poliploides ($\sim 12x$) y aneuploides con ~ 120 cromosomas (Le Cunff *et al.*, 2008; Grivet *et al.*, 2004) (Figura 1.4).



Figura 1.4. Cultivar moderno de caña de azúcar, CG98-46 (Híbrido *Saccharum* spp.).

En Guatemala, el Programa de Variedades del Centro Guatemalteco de Investigación y Capacitación de la Caña de Azúcar (CENGICANÑA) ejecuta su programa de mejoramiento genético con el principal objetivo de contribuir al incremento de la productividad de azúcar en la *Agroindustria Azucarera de Guatemala*, con la generación, selección o adaptación de variedades de caña de azúcar con producciones de caña y rendimientos de azúcar apropiados. En términos generales, estos objetivos se alcanzan mediante el estudio de la adaptación de las variedades de caña de azúcar en diferentes condiciones ambientales, así como con la determinación de la resistencia o tolerancia aceptable a las enfermedades y plagas más comunes (Orozco *et al.*, 2012). Por otro lado, se busca que las características de los genotipos se adapten al manejo agronómico convencional de la agroindustria azucarera guatemalteca, así como a los procesos de extracción de sacarosa y otros subproductos industriales de la caña de azúcar (Orozco *et al.*, 2012).

En caña de azúcar, como en muchos otros cultivos, la selección de los individuos de mejor desempeño es la clave para el mejoramiento general de la especie (Allard, 1980; Fehr, 1987). En este sentido, la selección de individuos superiores implica, invariablemente, la selección de alelos (genes) que mejor representen las características que se desean expresar. Esta selección de genes puede realizarse fenotípicamente (Fisher, 1936; Cerón-Rojas *et al.*, 2006),

genotípicamente (Lande y Thompson, 1990; Cerón-Rojas *et al.*, 2008a), o bien mediante la combinación de ambos.

El mejoramiento genético en caña de azúcar, una especie poliploide y aneuploide, es relativamente complicado en términos convencionales debido al complejo problema que supone el multialelismo (Singh *et al.*, 2010; Da Silva *et al.*, 1993). En ese sentido, los procedimientos generales de mejora genética en caña de azúcar se guían por la generación inicial de variabilidad genética y, posteriormente, la selección subsecuente, en diferentes estados, de genotipos fijados por clonación con el mejor desempeño desde el punto de vista agronómico e industrial (Orozco *et al.*, 2012).

En CENGICANÁ, Guatemala, el Programa de Variedades desarrolla sus actividades desde hace 20 años. La selección de clones y el desarrollo de variedades de caña de azúcar se realiza siguiendo patrones de selección basados en la medición de variables que permiten establecer las calidades fenotípicas de las selecciones, y sus cualidades, para obtener una variedad de uso comercial en el área de producción. Guatemala cultiva alrededor de 230,000 ha en más de 20 zonas agroecológicas (Orozco y Buc, 2016), lo que obliga al desarrollo de diferentes variedades que puedan adaptarse a diferentes zonas de producción, lo cual impone, además, la necesidad de realizar un mejoramiento genético eficiente.

Dentro de los limitantes más importantes en el mejoramiento genético de la caña de azúcar se encuentra la estrecha variabilidad genética. Los cultivares modernos de caña de azúcar derivan de cruza simples entre muy pocas variedades generadas hace unos 120 años (Rosales-Longo, 2011). Debido a este problema es necesario generar grandes poblaciones de individuos (de 160,000 a 200,000), mediante cruza simples o policruzas. Usualmente, de una generación o serie de cruzamientos, una variedad es seleccionada, y alcanza su cultivo comercial después de doce años desde su generación hasta su liberación. Además, se necesitan unos 5 años de desarrollo hasta su adopción y cultivo comercial. Por tanto, la selección de esa variedad es la clave del éxito de un programa de mejoramiento. Pero, ¿cómo saber seleccionar una buena variedad?

1.3 EL PROGRAMA DE VARIEDADES DE CENGICAÑA

En un proceso general de investigación en el campo del mejoramiento genético, se genera una línea general de investigación que establece, con orden, los procesos y metodologías a seguir para la buena obtención de los resultados buscados. En el caso del Programa de Variedades de CENGICAÑA, esta línea general de trabajo está establecida por el esquema general de investigación y mejoramiento genético (Figura 1.5). La cantidad de trabajo involucrado, desde el manejo del germoplasma, cruzamientos, selección y desarrollo comercial de las variedades, es muy amplio, por lo que integra a seis fitomejoradores. Entre los trabajos de investigación desarrollados se encuentran la optimización de los procesos de manejo del germoplasma, en forma convencional y mediante cultivo de tejidos; los estudios de la floración del cultivo, de la fisiología y de la respuesta a las modificaciones del fotoperiodo; así como otros que puedan ayudar a optimizar la recombinación genética (Quemé *et al.*, 2016).

En el apartado de la selección, el programa busca también mejorar sus procedimientos, con el fin de hacer lo más eficiente posible la obtención de nuevos genotipos, mediante la optimización de sus metodologías de diseño, evaluación, y selección. Lo anterior derivará en la obtención de nuevos genotipos de caña de azúcar de alto nivel de desempeño y, a más largo plazo, contribuirá en la elevación de los niveles de recuperación de sacarosa en la Agroindustria Azucarera Guatemalteca.

Para realizar la selección y desarrollo de una variedad, el Programa de Variedades de CENGICAÑA cuenta con varios pasos o estados de selección. En los diferentes estados de selección del Programa de Variedades de CENGICAÑA (Figura 1.5) se evalúan anualmente cerca de 160,000 diferentes genotipos.

Orozco *et al.* (2012) presentan al Programa de Variedades de CENGICAÑA según un proceso secuencial (Figura 1.5). Este proceso inicia con el mantenimiento y uso del recurso genético, esto incluye la conservación y la selección de progenitores para el siguiente paso o sub-proceso, el cual lo constituye el programa de cruzamientos, que es donde principalmente se obtiene todo el material genético recombinante para los siguientes sub-procesos. Sigue luego el programa de selección, el cual consta de cinco estados. El Estado I, es el primer sub-proceso de selección, donde se eligen cerca de nueve mil genotipos de entre más de 150,000.

Estas selecciones pasan al Estado II, todos los genotipos en este estado se evalúan en dos ambientes contrastantes de la zona de producción de la agroindustria azucarera de Guatemala, en experimentos con diseños aumentados.

El siguiente paso es el Estado III; en este estado se evalúan las selecciones hechas en el Estado II, las cuales son cerca de mil genotipos por año, en experimentos en diseños de bloques incompletos. Las selecciones del Estado III, que son entre 60 y 80 variedades, pasan a ser evaluadas en un grupo de experimentos mucho más ampliados, tratando de establecer los niveles de interacción de Genotipo por el Ambiente. Estos experimentos son denominados Pruebas Regionales o Estado IV, y se realizan en diseños de bloques completos al azar. Las selecciones del Estado IV pasan a ser evaluadas en el último estado de evaluación: El Estado V o las Pruebas Semi-comerciales. Las selecciones finales del Estado V, unas dos variedades, son finalmente tomadas para impulsar su adopción y desarrollo en una última fase del programa que se denomina “Apoyo al desarrollo comercial de las variedades liberadas”.

El presente estudio se circunscribe en el Estado III. En este estado los genotipos se evalúan, en ensayos en un diseño de bloques incompletos (alfa-látice), las características y el desempeño de alrededor de 1,000 genotipos de caña de azúcar (Rosales-Longo *et al.*, 2016). En cada ciclo de evaluación (período de zafra) se cuenta con un total de ocho ensayos evaluados en dos ciclos de producción (Plantía y Primera soca). En el Estado III usualmente se realizan unas 54 mediciones para el cálculo y establecimiento de 14 variables de estudio. Si bien se realizan análisis estadísticos, como análisis de variación para modelos convencionales (Montgomery, 1991) y mixtos (Bates *et al.*, 2015), la selección de los cultivares finalmente se realiza mediante el uso de estrategias basadas en la competencia del fitomejorador, lo que reporta un elevado grado de eficacia; sin embargo, consume tiempo y recursos. Lo anterior puede incidir en un proceso de selección largo y, posiblemente, con cierto margen de imprecisión, lo que supone un problema en el mejoramiento genético de la caña de azúcar.

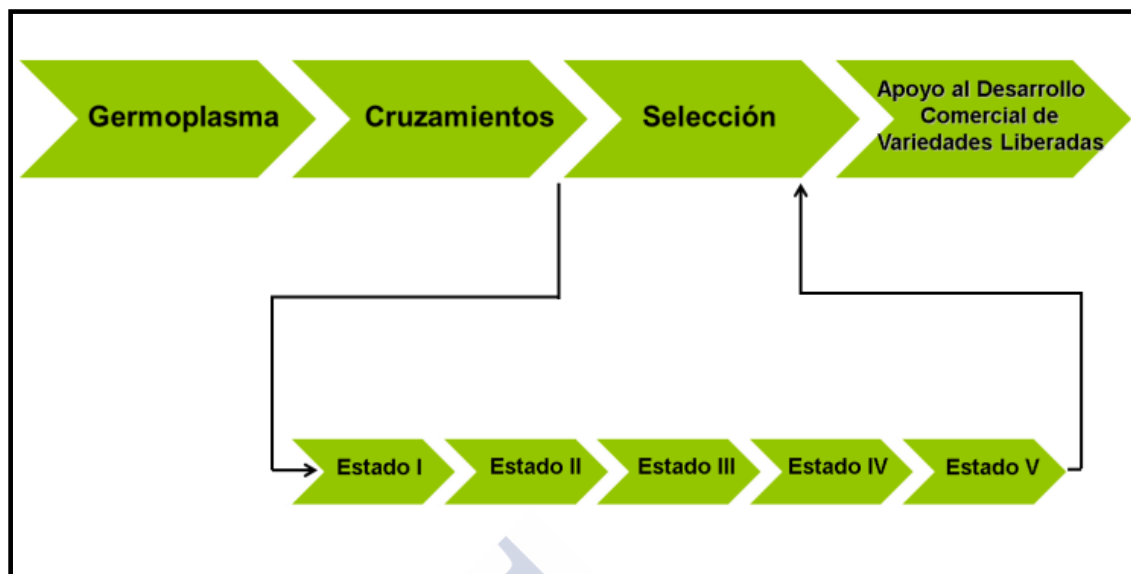


Figura 1.5. Esquema de trabajo para el mejoramiento de caña de azúcar en el Programa de Variedades de CENGICAÑA.

1.4 MEJORAMIENTO GENÉTICO DE LA CAÑA DE AZÚCAR

La caña de azúcar moderna (híbridos *Saccharum* spp.) (Figura 1.3) es genéticamente una especie muy compleja, por lo que su mejoramiento genético en el ámbito tradicional de la endocría e hibridación es problemático (Singh *et al.*, 2010; Da Silva *et al.*, 1993). Hoy en día, los cultivares modernos han reemplazado a los cultivares tradicionales de *S. officinarum* y a los clones de *S. spontaneum* (Grivet *et al.*, 2004).

Como producto de la inestabilidad genética, y debido a la presencia de muchas formas alélicas de los genes (debida ésta, a su vez, a la poliploidía), la mejora genética de la caña de azúcar no puede encararse desde el punto de vista del mejoramiento convencional, esto es: formación de líneas puras, hibridación y retrocruzas para la incorporación de características monogénicas (Fehr, 1987).

Por otro lado, la diversidad genética de los cultivares modernos de caña es estrecha y, en consecuencia, poco explotable, particularmente en cuanto al contenido de azúcar (Singh *et al.*, 2010; Babu *et al.*, 2006; Piperidis *et al.*, 2010; Ming *et al.*, 2006). Los cultivares modernos son derivados principalmente de no más de 20 genotipos de cultivares *nobilizados* (Ming *et al.*, 2006). La mayoría de programas de mejoramiento de la caña de azúcar en el mundo continúan usando esta base genética, buscando el cruzamiento de individuos pertenecientes a grupos genéticamente más o menos contrastantes, de tal forma que se pueda explotar la mayor

variabilidad posible (Quemé *et al.*, 2005). Actualmente se busca ampliar la base genética mediante el uso de aquellos clones usados originalmente en Java e India, producto de la *nobilización* original, este procedimiento está siendo continuado especialmente en Barbados (Ming *et al.*, 2006); sin embargo, los resultados de estos procedimientos aún no son evidentes.

En Guatemala, el mejoramiento genético convencional de la caña de azúcar sigue los siguientes procesos generales: a) selección de progenitores; b) generación de la variabilidad genética por la recombinación genética (cruzas entre genotipos modernos, cruzas con parientes silvestres o mediante la introducción de germoplasma); y, c) dentro de esta variabilidad se ejecuta un programa de selección de genotipos (Orozco *et al.*, 2012). Este procedimiento de trabajo es más o menos el mismo en diferentes programas de mejora genética de la caña alrededor del mundo (Berding *et al.*, 2004; Ming *et al.*, 2006; Piepho *et al.*, 2008).

La estrecha base genética de la caña de azúcar obliga a que la selección de genotipos en una población sea un asunto complicado. En el proceso de la realización de los cruzamientos, debido a que no es fácil desarrollar una recombinación que produzca una alta cantidad de buenos genotipos, se hace necesario realizar una gran cantidad de éstos, con el fin de encontrar aquellos que tengan condiciones para convertirse en variedades de uso comercial (Ming *et al.*, 2006; Orozco *et al.*, 2012). En el Programa de Variedades de CENGICANA, la proporción de individuos seleccionados según los estados de selección es muy baja, y va del 2% en el Estado I y II (Rosales-Longo *et al.* 2012a;) al 5-13% en el estado III (Rosales-Longo *et al.*, 2016b). En términos generales, lo importante en un programa de mejoramiento genético es la ganancia genética. Se busca que, con cada ciclo de selección, los genotipos seleccionados sean cada vez más productivos y con buenas características, adecuadas para los usuarios (Allard, 1980). Es en el proceso de selección donde la ganancia genética puede ser medida (Fehr, 1987).

1.4.1 La interacción Genotipo \times Ambiente ($G \times E$)

En general, los fitomejoradores están obligados a lidiar con caracteres de orden cuantitativo (Smith, 1936). La dificultad con este tipo de caracteres es que están altamente influenciados por el ambiente. La gran mayoría de los caracteres de importancia tienen un tipo de variación cuantitativa (Bos y Caligari, 2007). El principal objetivo del fitomejorador es justamente el mejoramiento, por medio de la selección, del valor genotípico concerniente a dichos caracteres. Para el fitomejorador es de mucha importancia conocer y entender cuál es el grado en que la expresión fenotípica de los caracteres con variación cuantitativa es debida a su conformación genotípica.

La variación cuantitativa es debida a dos principales factores (Bos y Caligari, 2007):

- a. **La variación en la calidad de las condiciones de cultivo.** El genotipo tiene una influencia parcial en la variación fenotípica. La variación en las condiciones de cultivo induce variación en la expresión fenotípica. El tamaño de la variación fenotípica en poblaciones genéticamente homogéneas refleja el balance entre la fortaleza del control genético en la expresión y el tamaño del efecto de la variación en las condiciones de cultivo.
- b. **La variación genética.** La expresión de características con variación cuantitativa puede ser afectada por un gran número de *loci*. Con un antecedente genético común, diferentes genotipos con un solo *locus* diferente pueden mostrar pequeñas diferencias en su expresión; pero estas diferencias en la expresión pueden ser mayores si los genotipos acumulan un mayor número de variaciones cuantitativas, es decir un mayor número de genes. En el caso de la caña de azúcar esto puede ser un mayor número de alelos (Ming *et al.*, 2006; Jackson 2005). El mismo caso ocurre con los caracteres asociados a la concentración de azúcar en caña de azúcar, los cuales, además, muestran altos valores de heredabilidad (Aitken *et al.*, 2006; Jackson, 2005; Inman-Bamber *et al.*, 2011).

En caña de azúcar, los caracteres de mayor interés, como la producción de biomasa y la concentración de azúcar, son caracteres poligénicos y, en consecuencia, altamente influenciados por el ambiente y, por tanto, con una alta variación cuantitativa (Berding *et al.*, 2004; Ming *et al.*, 2006). Los caracteres como el rendimiento de caña son usualmente continuos en su expresión bajo condiciones ambientales normales de producción (Skinner *et*

al., 1987). Algunos aspectos sobre la variación cuantitativa y sus bases genéticas se presentan en la Tabla 1.3, con las ecuaciones correspondientes para la estimación de la heredabilidad (Skinner *et al.*, 1987; Holland *et al.*, 2003). Este tema es relevante porque, según los valores de heredabilidad de los caracteres importantes, se logra apreciar el nivel de influencia del ambiente. En general, la heredabilidad es la proporción de la variabilidad de un carácter en particular atribuible al genotipo, con respecto a la variabilidad fenotípica total (Fehr, 1987).

Tabla 1.3. Aspectos sobre variación cuantitativa para especies de reproducción sexual y vegetativa.

	Reproducción vegetativa	Reproducción sexual
Fenotipo	$p = g + e$	$p = a + d + i + e$ ($a+d+i=g$)
Heredabilidad	$h^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$ <i>Heredabilidad en el sentido amplio.</i>	$h^2 = \frac{\sigma_a^2}{\sigma_p^2}$ <i>Heredabilidad en el sentido estrecho.</i>

p =fenotipo, g = valor genotípico, e =desviación ambiental, a =valor de mejora genética (aditivo), d =desviación por la dominancia, i =desviación por la interacción (epistasia), σ^2 =varianza.

Las condiciones ambientales afectarán mucho más a la expresión fenotípica cuando se evalúan en los estados iniciales de selección, como en la selección por familias por ejemplo, donde las unidades experimentales son pequeñas y los valores de heredabilidad (sentido amplio) son relativamente bajos para los caracteres de importancia económica, como la producción de caña por hectárea, el número, la altura y el diámetro de los tallos, y otras. En tanto que, las medidas directas del contenido de sólidos solubles (°Brix) y las mediciones de resistencia a enfermedades presentan valores más altos de heredabilidad (Skinner, *et al.*, 1987; Silva *et al.*, 2011).

En experimentos de evaluación de genotipos de caña de azúcar, en múltiples ambientes y en varios ciclos de evaluación (años), se observa que la influencia del ambiente es mayor a la que presentan, por sí solos, los genotipos y las localidades (Berding *et al.*, 2004; Quemé *et al.*, 2007; Kimbeng *et al.*, 2009; Rodríguez *et al.*, 2013). Quemé *et al.* (2007) encontraron significancia estadística en todos los factores involucrados (localidades, ciclos de evaluación y genotipos) y, en menor grado, las interacciones correspondientes. Por otro lado, Jackson y Hogarth (1992) concluyen que no hay mayor ganancia en realizar muchas evaluaciones a

través de varios ciclos. En este sentido, de dos a tres años de evaluación deberían ser suficientes. Como consecuencia de lo anterior, evaluar una gran cantidad de sitios es mucho más informativo que las evaluaciones a través de los años, en especial en temas como la habilidad del rebrote de la caña en varios períodos (socas) de evaluación (Berding *et al.*, 2004).

1.4.2 Procedimientos de Selección

En caña de azúcar, el procedimiento de selección está fundamentado en los temas de interacción $G \times E$ (Skinner *et al.*, 1987; Jackson *et al.*, 1995; Quemé *et al.*, 2007; Quemé *et al.*, 2010; Rodríguez *et al.*, 2013). En este sentido, la selección de genotipos está establecida según la variación cuantitativa de los caracteres de importancia económica para la especie. Por lo anterior, la selección de clones no se realiza en un único, gran y concluyente experimento, como sería de suponer para genotipos fijados por clonación, sino en evaluaciones secuenciales, ya que la influencia del ambiente variará con respecto a las localidades y a las condiciones de los años de evaluación (Skinner *et al.*, 1987; Orozco *et al.*, 2012). En este tipo de experimentos, el número y tamaño de las unidades experimentales está más determinado por el número de genotipos y áreas disponibles que por el nivel predeterminado de probabilidad de las comparaciones estadísticas (Skinner *et al.*, 1987).

La selección de las variedades de desempeño superior puede hacerse, inicialmente, por simples observaciones generales de las características deseables para un genotipo en particular. Esta es una forma muy simple de selección y usualmente es muy eficiente, particularmente para aquellas características gobernadas por un solo gen (mongénicas) (Fehr, 1987; Allard, 1980). De forma similar a lo que realizó Mendel (1866), la observación de genes directamente relacionados con caracteres observables, como el color y forma de las semillas, en caña de azúcar la selección de los genotipos superiores también puede ser realizada por observación directa del fenotipo.

Sin embargo, como se apuntó anteriormente, la consecuencia en el campo de la alta ploidía de la caña de azúcar y, en consecuencia, la presencia de muchos alelos para un mismo gen, resulta en una alta interacción entre el genotipo y el ambiente en aquellos caracteres de importancia económica (Skinner *et al.*, 1987; Berding *et al.*, 2004; Rodríguez *et al.*, 2013; Salazar y Victoria, 2013). Por tanto, la simple observación no es suficiente para establecer el

grado de la ganancia genética en la selecciones. En consecuencia, es necesario establecer procedimientos que permitan desglosar de la mejor manera posible el efecto *genético* del efecto *ambiental*.

En este sentido, existen varios procedimientos, principalmente estadísticos, que permiten separar los efectos ambientales de los genéticos en un proceso de selección, tales como los análisis de variación (Montgomery, 1991), los Análisis de Componentes Principales (ACPs) en modelos aditivos (Quemé *et al.*, 2007; Quemé *et al.*, 2010; Bissessur *et al.*, 2010; Ramburan *et al.*, 2011), los modelos lineales mixtos (De Oliveira *et al.*, 2013; Piepho *et al.*, 2008; Pastina *et al.*, 2012; Taylor y Verbyla, 2011) y los modelos de análisis basados en Componentes Principales (CPs) como los propuestos por Cerón-Rojas *et al.*, (2006) y Cerón-Rojas *et al.*, (2008b).

Adicionalmente, los marcadores genéticos (ADN) también han permitido la caracterización de los genotipos de caña de azúcar (Ming *et al.*, 2001; Aitken *et al.*, 2006; Bernardo, 2008; Simon y Hemaprabha, 2010). Los marcadores de ADN son útiles para establecer diferencias entre clones, con marcadores monogénicos o poligénicos (QTL's). Incorporar estas variables en un procedimiento de selección puede ser importante porque agrega valor a los procedimientos de mejora genética en caña de azúcar mediante la adición de variables que no son afectadas por el ambiente. Muchas de las variables de importancia económica como la producción de azúcar, resistencia a enfermedades bacterianas y biomasa están gobernadas por la acción aditiva de varios o muchos genes (QTL's) (Semagn, *et al.*, 2010; Franco *et al.*, 2013). La formación de grupos de ligamiento y la asociación de características es muy útil también para asistir a la selección.

Debido en parte a todo lo descrito anteriormente, el Programa de Variedades de CENGICANA ejecuta su trabajo de mejora genética mediante evaluaciones secuenciales de los genotipos de caña de azúcar, generados o incorporados al programa, de tal manera que cada fase de evaluación (Estado) es un filtro que permite la observación de un gran número de genotipos (160,000) en los estados iniciales de selección (Estados I, II), un número intermedio de genotipos (1,000) en el Estado III y un número bajo de genotipos (25) en los estados avanzados (Estados IV y V) (Orozco *et al.*, 2012).

En la Figura 1.5 se presentó el esquema general de trabajo del Programa de Variedades de CENGICAÑA. En el subprograma de selección, con respecto al Estado II, anualmente, se evalúan entre 6,000 y 10,000 genotipos en dos localidades contrastantes de la zona de producción de caña de azúcar en Guatemala y dos patrones de floración diferentes (**flor y no_flor**) (Rosales-Longo *et al.*, 2016a), lo que deriva en la instalación de cuatro diferentes experimentos. Las evaluaciones son visuales, y las mediciones importantes se centran en la concentración de sólidos solubles (°Brix). Debido al gran número de genotipos se hace difícil hacer repeticiones, por lo que se opta por el uso de un diseño experimental aumentado (Cullis *et al.*, 2006; Federer y Raghavarao, 1975), donde se aumenta solamente el número de repeticiones de las variedades testigo. Los resultados arrojan entre 100 y 150 variedades seleccionadas por experimento. Seguidamente, se conducen cuatro experimentos para cuatro subprogramas de selección, es decir, finalmente se seleccionan entre 400 y 600 genotipos para ser evaluados en el Estado III. Los genotipos evaluados en el Estado II son seleccionados al finalizar el segundo ciclo (soca) de cultivo (Rosales-Longo *et al.*, 2016a). A partir de esta selección se cuenta con una mayor cantidad de material vegetal para evaluar los clones en un estado más avanzado de evaluación y selección, El Estado III.

En el Estado III de selección se cuenta con dos series de genotipos generados en años consecutivos, por lo que se obtienen entre 1,000 y 1,200 genotipos anualmente para su evaluación y selección (Rosales-Longo *et al.*, 2016b) de dos series diferentes, CG12 y CG13, por ejemplo. En este estado se plantan experimentos en un diseño de bloques incompletos parcialmente balanceados (Montgomery, 1991) y se inician las mediciones correspondientes para las estimaciones de los caracteres de importancia económica, como las toneladas de caña por hectárea, el rendimiento de azúcar por hectárea, el rendimiento de materia seca por hectárea, y otros caracteres. Como se aprecia, los principales caracteres en evaluación tienen variación cuantitativa, por lo que su tratamiento es desde el punto de vista de la variación continua y la interacción $G \times E$.

En el Estado III de selección se utilizan 14 variables de valoración para los genotipos candidatos, y la selección se hace a partir de las expresiones fenotípicas consideradas más importantes, como la producción de azúcar por hectárea y la producción de biomasa por ha (caña fresca y materia seca) (Rosales-Longo *et al.*, 2016b). Sin embargo, existe un importante número de estas 14 variables que tienen importancia relativa y que son útiles para la toma

final de decisiones, por lo que resulta de interés generar funciones que permitieran incluir a las 14 variables de valoración de cada genotipo en evaluación.

1.5 LOS ÍNDICES DE SELECCIÓN

Entre los factores limitantes en la mejora genética de la caña de azúcar, como ya se indicó anteriormente, se encuentran su estrecha variabilidad genética, la variación cuantitativa de los caracteres de interés, el alto número de genotipos en evaluación y otros.

En general el objetivo de un programa de mejora genética en caña de azúcar es la mejora en términos de (Berding *et al.*, 2004; Orozco *et al.*, 2012):

1. la productividad (Toneladas de Caña por hectárea -TCH- o Toneladas de Azúcar por Hectárea -TAH-),
2. la resistencia a enfermedades,
3. la resistencia a plagas, y
4. las buenas características de molienda y fábrica.

Según lo expuesto con anterioridad, el grado de precisión de la selección puede ser afectado por la dificultad de realizar una valoración detallada de cada genotipo en evaluación. Si fuera posible compilar en un solo valor todos los valores de las variables de caracterización para cada genotipo, con cada una de estas variables ponderadas, según su importancia, sería posible mejorar la evaluación y selección de estos genotipos.

Un índice de selección es un valor que permite la selección de genotipos mediante la valoración simultánea de varias características asociadas a una en particular (producción de biomasa por ejemplo), el cual permite diferenciar individuos (Smith, 1936). Si los valores de una característica correlacionan con otra característica, puede hacerse una selección con base a una de ellas, si la medición de una de las dos es complicada. Extendiendo el concepto un poco más: si el efecto aditivo de varias características correlacionan con una variable de interés, por ejemplo la producción de biomasa, entonces puede estimarse mediante una regresión el valor predicho de la producción de biomasa mediante la combinación de aquellas características. Si se extiende un poco más aún el concepto, y de muchas variables puede evaluarse su efecto aditivo, entonces éstas en combinación, en una regresión, producirán un

valor que en general puede ser un indicativo del valor genotípico total de un genotipo en evaluación, todo esto a partir de valores fenotípicos. Este último valor puede ser considerado un índice que permita diferenciar entre individuos (Bos y Caligari, 2007).

Smith (1936) definió el índice de selección como una combinación lineal de los valores fenotípicos observados de la expresión de la característica de interés; y son generalmente utilizados para discriminar dentro de las unidades de selección, considerando tanto la estructura genética como probabilística de la población de donde los genotipos fueron originados, así como la importancia económica de las características en estudio, ya que según Smith (1936), los valores de los componentes de una regresión son ponderados según su importancia económica. En consecuencia, solamente los genotipos de superior valor económico son reproducidos (Cerón-Rojas *et al.*, 2006).

De acuerdo con Cerón-Rojas *et al.* (2006), el principio de un índice de selección puede definirse en el siguiente contexto: buscar seleccionar a los materiales con altos valores genotípicos (g) (los cuales no son directamente observados), y que están relacionados con los valores fenotípicos observables o medibles (p). Se asume que los valores genotípicos no observables y sus valores fenotípicos observables tienen una probabilidad conjunta de distribución, tal que, si se realiza una regresión entre los valores genotípicos y fenotípicos, los valores (\hat{g}) de dicha regresión se consideran un buen estimador de g , dado el valor de p , esto es $E(g/p)$. De esta forma los genotipos con los valores más altos de dicha regresión, son los de mayor valor genético y, por tanto, pueden ser sujetos de selección. El cálculo de una función de regresión (índice de selección) brindará un buen estimador de los valores genéticos de los genotipos en evaluación.

En resumen, un índice de selección es calculado entonces con base en el desempeño de varias características medibles. En la metodología inicial de los índices de selección, los valores de los componentes de una regresión son frecuentemente ponderados según su importancia económica (Smith, 1936). Esto es particularmente difícil de seguir, por cuanto los aspectos económicos son complicados de usar como ponderadores, ya que su estimación requiere de estudios amplios, además de que pueden haber fluctuaciones en cuanto a precios en los mercados y, en consecuencia, cambios en los ponderadores económicos.

Los índices de desempeño o de selección son valores que ayudan a diferenciar genotipos, cuando las variables de valoración son de variación cuantitativa. En general, esto ayuda cuando las variables de más alto impacto económico son del orden cuantitativo, como es el caso de las variables de valoración en las evaluaciones de genotipos de caña de azúcar.

A continuación se presenta una relación de los índices de desempeño calculados y evaluados en el presente estudio.

1.5.1 Índice Función Discriminante Para la Selección de Plantas (Smith, 1936).

Un abordaje detallado sobre la función discriminante para la selección de plantas se encuentra en el trabajo de Smith (1936). En este apartado solamente se hace una breve reseña del fundamento teórico del índice.

Los caracteres de variación cuantitativa presentan una dificultad particular, puesto que las variaciones heredables son enmascaradas por variaciones no heredables que hacen difícil establecer el valor genotípico de los materiales en evaluación, a menos que se cuente con una gran cantidad de material de propagación que permita hacer una gran cantidad de parcelas replicadas para cada genotipo (Bos y Caligari, 2007).

Smith (1936) señala que, en los estados tempranos de selección, el fitomejorador intenta seleccionar plantas en el campo sobre la base de caracteres observables (componentes del rendimiento), los cuales él cree que están asociados con un carácter deseable (el rendimiento); sin embargo, el verdadero valor atribuible a cada carácter es, en realidad, desconocido. El problema puede ser abordado mediante la búsqueda de una función lineal discriminante (Fisher, 1936), en la cual los valores de los caracteres observables pueden ser un indicativo del “valor genético” de los genotipos. Los procedimientos sugeridos por Smith (1936) (Optimum Selection Index) tienen ventajas, como la de ponderar mayormente a aquellas variables que son de alta importancia, y es relativamente simple de analizar. Como desventajas tiene que requiere de una gran cantidad de información para establecer con precisión los ponderadores económicos, los errores de muestreo pueden ser importantes y los pesos económicos son difíciles de establecer (Cerón-Rojas, *et al.*, 2006).

Siguiendo la teoría discutida por Smith (1936) en un programa de selección de caña de azúcar, si se desea evaluar un conjunto de genotipos, principalmente con respecto a los

valores del producción general de biomasa, pureza de los jugos, resistencia a una enfermedad en particular, y así sucesivamente ($x_1, x_2, x_3, \dots, x_n$), debe asignarse un nivel de ponderación económica (b_n) a cada una de estas variables. Por ejemplo, el rendimiento de biomasa es el 100 por ciento de la importancia económica, por tanto su valor de ponderación es “1”; por otro lado, la pureza de los jugos está relacionada con su productividad en la fábrica, y supongamos que un valor menor al 80% de pureza infringe una reducción del 10% en la recuperación de sacarosa, y que una infección con una enfermedad en particular puede reducir la calidad de la materia prima y, en consecuencia, una reducción en un 20% de la producción; y así sucesivamente se pueden establecer los ponderadores económicos ($b_1, b_2, b_3, \dots, b_n$), donde b_n corresponden a las ponderaciones económicas según la importancia de la variable, donde x_n corresponde a las expresiones fenotípicas del material según la variable estudiada. Todos los valores de b_2, b_3, \dots, b_n , deben ser luego referidos al de importancia mayor, en este caso el rendimiento de biomasa, b_1 .

Los valores fenotípicos y sus ponderadores económicos, pueden representarse como sigue:

$$Y = b_1x_1 + b_2x_2 + \dots + b_nx_n \quad 1.1$$

donde, b_n = ponderadores económicos y x_n = fenotipos de las variables de valoración.

El valor genotípico de x puede denotarse como ξ , entonces el valor genotípico de una línea o clon en evaluación sería denotado por:

$$\psi = a_1\xi_1 + a_2\xi_2 + \dots + a_n\xi_n \quad 1.2$$

El principal problema de la Ecuación 1.2 es que los verdaderos valores genotípicos no pueden ser observados. En contraste se cuenta solamente con los valores fenotípicos, los cuales no representan exactamente a los valores genotípicos.

Por otro lado, el valor de genético cualquier observación de cada carácter se puede descomponer así:

$$x_i = \xi_i + \epsilon_i \quad 1.3$$

donde, el valor de cada fenotipo x_i es igual a la suma del efecto genotípico (ξ_i) y el efecto ambiental (ε_i). Además, se asume que estas dos partes son independientes, de tal forma que la covarianza entre ξ_i y ε_i tiene una esperanza de cero. Entonces sí, t_{ii} , g_{ii} y e_{ii} son las varianzas de x_i , ξ_i y ε_i respectivamente, $t_{ii} = g_{ii} + e_{ii}$.

El problema ahora será encontrar el valor de Y , el cual es el resultado de la acción aditiva de las variables descritas. Para esto Smith propone el uso del concepto de la *función discriminante* de Fisher (1936), a fin de hacer una integración de las áreas correspondientes a una probabilidad de selección (q) (Figura 1.6) a un valor específico de selección diferencial en unidades estándar (Y') (Smith, 1936). Los genotipos que igualen o superen el valor Y' serán sujetos de selección. La teoría estadística detrás de estos procedimientos es larga y compleja.

Uno de los principales escollos es la realización de todos los estudios para establecer la importancia económica de cada variable de valoración. Esta función discriminante requiere amplios estudios, y procedimientos estadísticos y económicos para establecer la proporción de importancia para cada variable según lo muestra Jackson *et al.* (2015).

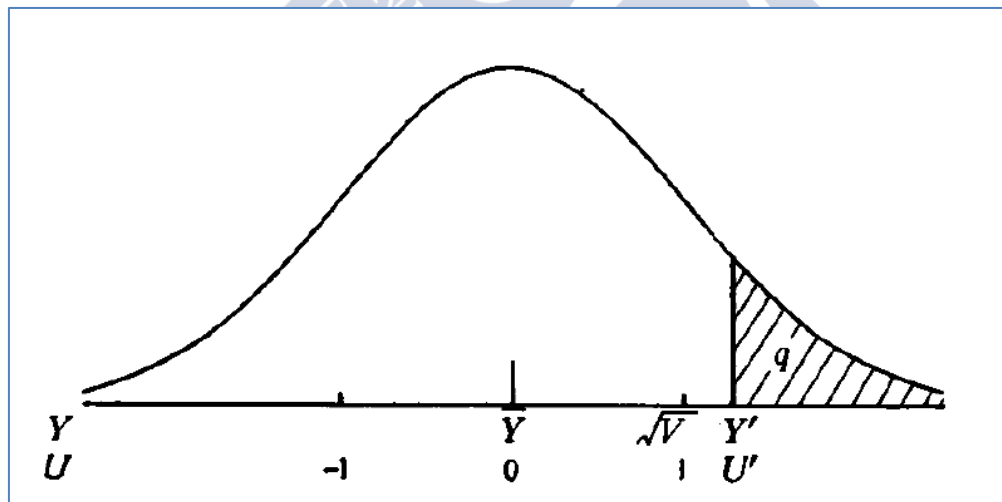


Figura 1.6. Frecuencias de distribución según la función discriminante de Fisher (Smith, 1936).

La idea general del índice de Smith es hacer una regresión entre los valores de las funciones de ψ y Y (Ecuaciones 1.1 y 1.2), de tal manera que los valores fenotípicos sean una estimación de los valores genotípicos. En este sentido, se hace necesario el cálculo de la

varianza de Y (V) ($V = E[Y - \bar{Y}]^2 = \sum_1^n \sum_j^n b_i b_j t_{ij}$) y la covarianza (W) entre ψ y Y [$W = E(\psi - \bar{\psi})(Y - \bar{Y}) = \sum_1^n \sum_j^n a_i b_j g_{ij}$].

Entonces, la regresión entre ψ y Y estará dada por:

$$B = \frac{W}{V} = \frac{\sum \sum a_i b_j g_{ij}}{\sum \sum b_i b_j t_{ij}} \quad 1.4$$

El valor más alto de la regresión será el correspondiente al índice de selección para cada observación. En este sentido, la maximización de la regresión B , mediante derivación es igual a:

$$\frac{\delta \ln B(V)^{\frac{1}{2}}}{\delta b_1} = \frac{\sum_j a_j g_{1j}}{W} - \frac{\sum_j b_j t_{1j}}{V} \quad 1.5$$

Esta maximización es el punto central de la propuesta de Smith (1936), y es la que formaliza la función discriminante para cada observación. Al despejar la Ecuación 1.5, el resultado se aplica a cada observación de la matriz de las “ n ” funciones discriminantes asociadas cada una, a cada genotipo u observación en evaluación. En su forma simplificada, por notación de matrices, la función discriminante es:

$$\sum_j b_j t_{1j} = \frac{V}{W} \sum_j a_j g_{1j} = K A_1 \quad 1.6$$

De esta forma se pueden obtener n ecuaciones de la forma general:

$$b_1 t_{11} + b_2 t_{12} + \cdots + b_n t_{1n} = K A_1;$$

$$b_1 t_{12} + b_2 t_{22} + \cdots + b_n t_{2n} = K A_2;$$

.....,

$$b_1 t_{1n} + b_2 t_{2n} + \cdots + b_n t_{nn} = K A_n$$

Estas expresiones pueden ser escritas brevemente como $\sum_{j=1}^n b_j t_{ij} = K A_i$; donde $K A_i$ es la representación matricial de las n funciones discriminantes ($A_i = \sum_j a_j g_{ij}$, y, $K=V/W$, g = a la varianza genética); e i es constante en cada ecuación.

Las soluciones de las n ecuaciones están dadas por:

$$b_j = K \sum_i A_i c_{ij} \quad 1.7$$

Debido a que el interés es solamente en los valores relativos de b , no es necesario evaluar K , la cual es constante en todas las ecuaciones, por lo que es suficiente evaluar solamente $b_j | K = \sum_i A_i c_{ij} = d_j$.

En resumen, la regresión entre los valores genotípicos y fenotípicos está dada por:

$$B(V)^{1/2} = \sqrt{\sum_i d_i A_i} \quad 1.8$$

Y finalmente, el valor genético o la “Respuesta a la selección” (R) de los materiales en evaluación, por sustitución en la ecuación 1.4 es,

$$\left(\frac{z}{q}\right) \sqrt{\sum d A} \quad 1.9$$

Donde z/q es una constante dependiente de la intensidad de selección q (Figura 1.6), la cual viene dada en unidades estándar. La dimensional que arroja la Ecuación 1.5 está dada por la correspondiente a la variable de referencia inicial (x_I), cuando se habla de componentes del rendimiento; en un caso contrario, la dimensional está dada por el valor más alto de b en cualquiera de los caracteres x_n .

De acuerdo con Smith (1936), todo el proceso de cálculo se fundamenta en tres premisas hipotéticas: *a)* Los efectos genotípicos y ambientales son aditivos para la magnitud de un carácter; *b)* Estos dos efectos son independientes, es decir su covarianza es igual a cero; *c)* Y y ψ están normalmente distribuidas.

1.5.2 Índices de selección por restricción de variables (*Restricted Selection Index*). (Kempthorne y Nordskog, 1959)

El detalle teórico y de uso práctico sobre este procedimiento es presentado por Kempthorne y Nordskog (1959). No es la intención de esta sección hacer un tratado extenso de los fundamentos teóricos de este índice, sin embargo, se presentan algunos aspectos que pueden ser de importancia sobre este procedimiento.

De acuerdo con Kempthorne y Nordskog (1959), el procedimiento presentado por Smith (1936), y que fue discutido en el epígrafe inmediato anterior, es aplicable solamente para cultivos autopolinizados y, esencialmente, no involucra teoría genética. En general, siguiendo las premisas teóricas de Smith (1936), Kempthorne y Nordskog (1959) plantean:

1. El fenotipo es igual a la suma de los efectos genéticos y ambientales ($P = G_i + E_i$). Las interacciones $G \times E$ son adicionadas al efecto ambiental.
2. El valor genotípico es enteramente asignado al efecto aditivo de todos los genes involucrados en el carácter de variación cuantitativa, también llamado valor de mejora genética (VMG) o “*breeding value*” en inglés.
3. El valor económico genotípico es denotado por ($H = \sum a_i G_i$), donde a_i son las constantes económicas y G_i son los valores genotípicos para cada una de las “ m ” variables o atributos.
4. Una regresión de H con respecto a P_i es lineal.

La segunda condición raramente ocurre, y se estima mediante los mínimos cuadrados del desempeño de los genotipos en un modelo matemático que incluya la evaluación de los efectos genotípicos y ambientales. Para Kempthorne y Nordskog (1959) el valor genotípico (G_i) es lo mismo que el “*breeding value*”, considerando que este último es una aproximación del primero.

Sobre las premisas indicadas, la selección se realiza con base en la función lineal siguiente, denotada como Restricted Selection Index, “ T ”.

$$I = \sum_i b_i P_i \quad 1.10$$

donde, \mathbf{b}_i corresponde al vector de los valores o ponderadores económicos de cada valor genotípico ($\mathbf{a}_i\mathbf{G}_i$). La ganancia genética en \mathbf{H} es igual a la correlación entre \mathbf{H} e \mathbf{I} , multiplicada por el diferencial de selección ($\beta_{HI}\Delta_I$); escrita de otra forma:

$$\frac{\text{Cov}(\mathbf{H}, \mathbf{I})}{\sigma_I^2} \Delta_I = \rho_{HI} \sigma_H \left(\frac{\Delta_I}{\sigma_I} \right) \quad 1.11$$

Δ_I es la diferencia de la media de \mathbf{I} después de la selección con respecto a la media antes de la selección, es decir, el diferencial de selección.

La Ecuación 1.11 muestra que el cambio genético debido a la selección es proporcional a ρ_{HI} y a $\left(\frac{\Delta_I}{\sigma_I}\right)$, el cual es el diferencial de selección en el índice, en unidades estándar. Si \mathbf{I} es distribuido normalmente, y los valores de ponderación (\mathbf{b}_i) son conocidos y sin error, y la selección se realiza en una presión porcentual específica “ p ”, entonces la relación de “ p ” con respecto $\left(\frac{\Delta_I}{\sigma_I}\right)$ es dada por ecuaciones de integración del área bajo la curva normal, y que establece los valores mínimos para el índice de los genotipos sujetos de selección (Ecuación 1.12). Esta ecuación es detallada tanto por Smith (1936) como por Kempthorne y Nordskog (1959).

$$\left(\frac{z}{p} \right) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi p}} \quad 1.12$$

En general, la teoría de los índices de selección hasta aquí es la misma que presenta Smith (1936); el cambio se presenta cuando Kempthorne y Nordskog (1959) indican que es posible que el fitomejorador desee mantener sin cambios un número específico “ r ” de atributos, de tal forma que el número de atributos variables sean $m-r$. Esto significa que pueden dejarse sin modificación algunos atributos (r), es decir, “restringir” el número de atributos a ser considerados en el resultado final del índice según la siguiente notación:

$$\mathbf{V}_1 = \sum_{j=1}^m \mathbf{c}_{j1} \mathbf{G}_j; \quad \mathbf{V}_2 = \sum_{j=1}^m \mathbf{c}_{j2} \mathbf{G}_j; \quad \dots \quad \mathbf{V}_r = \sum_{j=1}^m \mathbf{c}_{jr} \mathbf{G}_j \quad 1.13$$

De esta forma la función lineal de discriminación es la misma $\mathbf{I} = \sum_i \mathbf{b}_i \mathbf{P}_i$, pero con las condicionantes siguientes: $\text{Cov}(\mathbf{I}, \mathbf{V}_k) = \mathbf{0}$ y que $k = 1, 2, \dots, r$, es decir, solamente con el

uso de las variables de interés. De esta forma, la correlación entre el valor de I y H , es decir, la función lineal y el valor genotípico están dados por:

$$\rho_{IH}^2 = \frac{\text{Cov}^2(I, H)}{\sigma_H^2 \sigma_I^2} = \frac{(\sum_{ij} a_i b_j G_{ij})^2}{(\sum_{ij} a_i a_j G_{ij})(\sum_{ij} b_i b_j P_{ij})} \quad 1.14$$

Todos estos cálculos pueden simplificarse al utilizar notación matricial (no mostrada aquí). En las operaciones de solución de las matrices, y debido a que los valores “ b ” no son afectados al ser multiplicados por una constante, se obtiene que $\sum_{ij} b_i b_j P_{ij} = b' P b = 1$, entonces la correlación en la ecuación originalmente mostrada en 1.11 se convierte en:

$$\rho_{HI}^2 = \frac{(a' G b)^2}{(a' G a)} \quad 1.15$$

En donde se satisfacen las condiciones de $\text{Cov}(I, V_k) = b' G c_k = c'_k G b$; $k = 1, 2, \dots, r$.

Todos los resultados obtenidos son necesarios maximizarlos siempre y cuando se cumplan las condiciones de restricción “ r ”: $\text{Cov}(I, V_k) = 0$ y $b' P b = 1$. Después del proceso de maximización y simplificación de ecuaciones, la ecuación del índice en su forma final queda como sigue:

$$b = a' G [I - P^{-1} G C (C' G P^{-1} G C)^{-1} C' G] P^{-1} G a \quad 1.16$$

La correlación entre I y H queda finalmente definida como:

$$\rho_{HI}^2 = \frac{a' G [I - P^{-1} G C (C' G P^{-1} G C)^{-1} C' G] P^{-1} G a}{a' G a} \quad 1.17$$

Los resultados de la utilización del índice es como sigue: el cambio en G_i , como resultado un cambio en Δ_I en el índice, es igual a:

$$\beta_{G_i I} \Delta_I = \frac{\text{Cov}(G_i, I)}{\sigma_I^2} \Delta_I = \frac{\text{Cov}(G_i, I)}{\sigma_i} \frac{\Delta_I}{\sigma_I}$$

1.18

Lo que es igual a:

$$\frac{\sum_j b_j G_{ij} \Delta_I}{\sigma_I \sigma_I}$$

1.19

La Ecuación 1.19 puede ser empleada en cualquier caso particular (Kempthorne y Nordskog, 1959).

1.5.3 Las Componentes Principales (CPs) como índices de desempeño

El Análisis de Componentes Principales (ACP) (Hernández, 1998; Jolliffe, 2002) trata de explicar la estructura de las variancias y covariancias de un conjunto de datos de variables x_i mediante unas cuantas combinaciones lineales de estas variables x_i . Estas combinaciones lineales son llamadas Componentes Principales (CPs), y no están correlacionadas entre sí; además, cada CP maximiza su variancia. EL ACP aspira a *reducir* o *simplificar* los datos y facilitar su análisis e interpretación (Hernández, 1998). Esta técnica es uno de los métodos de ordenación más comúnmente utilizados para proyectar el hiperespacio de muchas variables a un espacio más reducido de dos o tres dimensiones (Huntley, 2011).

El ACP calcula nuevas variables “hipotéticas”, llamadas **Componentes Principales** (CPs), como resultado de la combinación de las variables originales. En consecuencia, a cada observación se le asigna un “score” (CP score), el cual es el producto de modelar cada CP con respecto a los valores fenotípicos observados. Las CPs son adimensionales, sin embargo, pueden ser relacionados con las variables originales (Huntley, 2011).

Los coeficientes de las funciones lineales de las CPs, según cada variable original, son llamados *eigenvectors*, *loadings*, *cargas* o *vectores característicos*. La magnitud de cada uno de estos coeficientes indicará el nivel en el que cada variable original está asociada a una CP en particular. Cada CP será definida como alguna combinación de todas las variables originales a través de una simple ecuación. Estas ecuaciones son utilizadas para el cálculo de cada “score” de la CP para cada observación. Por otro lado, la ordenación de datos, producto del ACP, producirá una medida de la varianza para cada componente principal, también llamada *eigenvalue*, *latent root* o *raíz característica* (Hernández, 1998; Huntley, 2011). La

suma de los *eigenvalues* es equivalente a la varianza total de todas las variables originales. La proporción dada por la razón entre el valor del *eigenvalue* de una CP en particular con la varianza total determina la cantidad de la varianza explicada por esa CP (Jolliffe, 2002).

El ACP identifica la dirección de la máxima varianza en el hiperespacio multidimensional del conjunto de datos que sea sujeto de análisis (Huntley, 2011). Jolliffe (2002) plantea que dicha maximización sigue un procedimiento general como sigue:

sea $k = 1, 2, \dots, p$ el número de CPs y p es el número de variables (x_i) originales, de tal suerte que el número de CPs es $k = p$. La k -ésima CP está dada, en notación matricial, por $\mathbf{z}_k = \mathbf{a}'_k \mathbf{x}$, donde \mathbf{a}_k es un *eigenvector* de la matriz de varianzas/covarianzas correspondiente al *eigenvalue*, λ_k , con el valor más alto. Aún más, si \mathbf{a}_k es elegido para contener la unidad de longitud ($\mathbf{a}'_k \mathbf{a}_k = 1$), entonces la varianza de $\mathbf{z}_k = \lambda_k$ (Jolliffe, 2002).

La maximización de la varianza no se alcanza para un \mathbf{a}_1 , por lo que es necesaria la normalización por una restricción. La restricción impuesta es $\mathbf{a}'_1 \mathbf{a}_1 = 1$, es decir que la suma de los cuadrados de los elementos de \mathbf{a}_1 es igual a uno (Jolliffe, 2002). Mediante la técnica de los multiplicadores de *Lagrange* (Jolliffe, 2002), la maximización, también en notación matricial, se obtiene finalmente por:

$$(\mathbf{\Sigma} - \lambda \mathbf{I}_p) \mathbf{a}_1 = \mathbf{0} \quad 1.20$$

Donde \mathbf{I}_p es la matriz identidad ($p \times p$). De esta manera λ es un *eigenvalue* de la matriz de varianzas/covarianzas y \mathbf{a}_1 es el correspondiente *eigenvector*. Para decidir cuál de los p *eigenvectors* entrega la $\mathbf{a}'_1 \mathbf{x}$, con la máxima varianza, la cantidad debe ser maximizada por:

$$\mathbf{a}'_1 \mathbf{\Sigma} \mathbf{a}_1 = \mathbf{a}'_1 \lambda \mathbf{a}_1 = \lambda \mathbf{a}'_1 \mathbf{a}_1 = \lambda \quad 1.21$$

De esta manera, λ debe ser lo más grande posible. Por tanto, \mathbf{a}_1 es el *eigenvector* correspondiente al mayor *eigenvalue* de la matriz de varianzas covarianzas ($\mathbf{\Sigma}$), y la varianza de $\mathbf{a}'_1 \mathbf{x} = \mathbf{a}'_1 \mathbf{\Sigma} \mathbf{a}_1 = \lambda_1$ el más grande *eigenvalue*. Detalles sobre la teoría subyacente, y el cálculo algebraico de las CPs, se encuentran en Jolliffe (2002) y Hernández (1998).

La interpretación de un ACP funciona mejor cuando el conjunto de datos muestra una distribución multidimensional normal. Así mismo, para que el ACP sea útil, debe haber algún tipo de correlación entre las variables originales (Huntley, 2011). Si las variables se miden en

la misma unidad de medida, entonces el ACP se ejecuta sobre la matriz de varianzas/covarianzas. Si las dimensionales son en unidades diferentes para las variables involucradas, entonces es preciso estandarizar las variables, y luego utilizar la matriz de varianzas/covarianzas.

Como ya se había indicado, el ACP identifica la dirección de la máxima varianza en el hiperespacio multidimensional del conjunto de datos sujeto de análisis. Una descripción visual común de los ACP plantea que, una vez el eje de mayor varianza se ha identificado, el conjunto de datos se someten a una rotación ortogonal donde las distancias entre los puntos de los datos en el hiperespacio (espacio multidimensional) son preservadas. El eje de mayor varianza es conocido como la Primera Componente Principal (CP1). CPs adicionales son identificadas en el proceso de ordenación, y son todas ubicadas a 90° una de la otra en el hiperespacio. Cada CP subsecuente contabiliza para la mayoría de la varianza, sin contar con la guardada por las CPs previas (Huntley, 2011). La esperanza en un conjunto de datos altamente estructurado por un ACP es que la mayor parte de la varianza sea contabilizada (explicada) por unos pocos ejes.

Frecuentemente, los primeros k de las CPs, con k mucho menor que p , explican una alta proporción de la variabilidad total de la x_i variables originales. Por lo tanto, estas CPs pueden representar a las p variables originales sin pérdida importante de información (Hernández, 1998).

En resumen, la contribución o importancia (*eigenvector*) de la i -ésima variable a la k -ésima CP se mide por el elemento \mathbf{a}'_i del vector normalizado \mathbf{a}'_i , asociado a la raíz característica (*eigenvalue*) λ_j de la estimación (\mathbf{S}) de Σ (matriz de varianzas/covarianzas) (Hernández, 1998). Al encontrar las CPs con los *eigenvalues* más altos, se contará con una buena proporción de la variabilidad explicada en una menor cantidad de funciones lineales (CPs) sin pérdida importante de información.

Adicionalmente, se conocen los *eigenvectors* para cada CP por lo que es posible identificar las variables que están asociadas con determinadas CPs. Por tanto, los *scores* (la función modelada según los valores reales de cada variable para cada observación) de las CPs son valores que resumen la combinación de las variables y, en consecuencia, pueden dar un único valor de caracterización a cada observación; es decir, *un índice*, en el cual se resume el

desempeño de la observación en la combinación total de todas las variables en estudio, según la función de determinada CP. De esta manera, dos o tres CPs pueden brindar valores que, en combinación, caracterizan a una observación, en lugar de las originales p variables. Este principio es el que se puede utilizar como un índice de desempeño.

1.5.4 Índice de selección basado en Eingenanalysis (*Eigenanalysis Selection Index Method, ESIM*) (Cerón-Rojas *et al.*, 2006)

Este procedimiento se fundamenta en la utilización de un ACP, en donde el primer *eigenvector* del *eigenvalue* más elevado, esto es, la primera CP, se utiliza como criterio de selección (Índice de Selección), y sus elementos determinan la proporción con la cual la característica asociada a ese *eigenvector* contribuye con el índice de selección.

Otra forma de representar la Ecuación 1.4, según Cerón-Rojas *et al.* (2006), es:

$$S\beta = \frac{\beta'S\beta}{\theta'\Sigma\beta} \Sigma\theta \quad 1.22$$

donde, $S\beta$ corresponde a la matriz de varianzas/covarianzas de la función descrita en la Ecuación 1.1; $\theta'\Sigma\beta$ es la covarianza entre las funciones fenotípica Y y genotípica ψ (Smith, 1936), o como las denotan Cerón-Rojas *et al.*, (2006): [$Cov(Y, Z)$]. En tanto que $\beta'S\beta$ es la varianza de Y (σ_Y^2) (Cerón-Rojas, *et al.*, 2006).

Con las bases establecidas por Smith (1936) y resumidas en la Ecuación 1.8, Cerón-Rojas y Sahagún-Castellanos (2005) sugieren que, mediante un ACP o *eigenanalysis*, se puede calcular la “respuesta a la selección (R)”. Si se asume que la covarianza entre Y y Z ($\theta'\Sigma\beta$) es igual a la varianza de Y ($\beta'S\beta$), entonces:

$$R = \sigma_Z \rho_{YZ} = k\sigma_Z \frac{\theta'\Sigma\beta}{\sqrt{\theta'\Sigma\beta}\sqrt{\beta'S\beta}} = k\sqrt{\beta'S\beta} \quad 1.23$$

Por lo anterior, la maximización de R sería equivalente a la maximización de $\beta'S\beta$, es decir, la varianza de Y (σ_Y^2). Todo lo anterior, en un esquema de ACP, sería equivalente a:

$$(S - \lambda I)\beta = 0 \quad 1.24$$

donde, β y λ son el *eigenvector* y el *eigenvalue* de S , respectivamente. En consecuencia, $\beta'S\beta = \lambda$ y $R = k\sqrt{\lambda}$ (ver Ecuación 1.20). En contraste a todo lo anterior, ESIM (Cerón-Rojas *et al.*, 2006) no asume que $\beta'S\beta = \theta'\Sigma\beta$, según Cerón-Rojas y Sahagún-Castellanos (2005). ESIM (Cerón-Rojas *et al.*, 2006) asume únicamente que la matriz de varianzas/covarianzas multiplicadas por el ponderador económico es igual al *eigenvector* de la matriz fenotípica de varianzas/covarianzas; es decir, $\Sigma\theta = \beta$. Esta asunción implica (bajo la restricción de $\beta'\beta = 1.0$), que $\frac{\beta'S\beta}{\theta'\Sigma\beta} = \lambda$ y, en consecuencia, la Ecuación 1.22 toma la forma $S\beta = \lambda\beta$, de la cual se derivó la Ecuación 1.24, en donde se mantiene que β y λ son el *eigenvector* y el *eigenvalue*, respectivamente.

Cerón-Rojas *et al.* (2006) indican que la respuesta a la selección (R), tanto en el procedimiento de Smith ($R_S = \frac{k}{\sqrt{\lambda}}$) como en ESIM ($R_{ESIM} = \frac{k}{\sqrt{\beta'S\beta}} = \frac{1}{\sqrt{\lambda}}$), ofrecen el mismo resultado.

1.5.4.1 La estimación de los *Eigenvalues*, *Eigenvectors*, ESIM y la Respuesta a la selección de ESIM (Cerón-Rojas *et al.*, 2006)

La varianza de Y (Y_{σ^2}), también denotada por $\beta'S\beta$ y $\beta'\beta$, puede ser derivada por $\beta'\beta = 1.0$ y $(S - \lambda I)\beta = 0$ en tanto que $(S - \lambda I)$ sea una matriz singular, esto es, el determinante $|S - \lambda I|$ debe ser una función de λ , que es igual a cero.

La ecuación $(S - \lambda I = 0)$ genera un conjunto de polinomios con q raíces $\lambda_1, \lambda_2, \dots, \lambda_q$, las cuales son las potenciales soluciones. Los estimadores de la “máxima verosimilitud” de los q *eigenvalues* y los q *eigenvectors* son $\hat{\lambda}_i'$ y $\hat{\beta}_i'$, respectivamente, de los cuales $|\hat{S} - \hat{\lambda}_i I| \hat{\beta}_i = 0$, $i = 1, 2, \dots, q$, donde \hat{S} es un estimador de S . Estos resultados permiten escribir la estimación de ESIM como

$$\hat{Y}_i' = \hat{\beta}_i' p \quad 1.25$$

En tanto que el estimador de la respuesta a la selección (R) de ESIM es $R_{ESIM} = \frac{k}{\sqrt{\lambda}}$, tal como se había indicado anteriormente.

Una descripción más detallada sobre las propiedades de los estimadores de los *eigenvalues*, los *eigenvectors* y los estimadores de ESIM y RESIM, así como sobre algunos ejemplos, se presentan en Cerón-Rojas *et al.* (2006) y Cerón-Rojas *et al.* (2008b).

1.5.5 Índice de selección restrictivo por medio de “Eigenanalysis” (Restricted Selection Index Method Based on Eigenanalysis, RESIM) (Cerón-Rojas, *et al.*, 2008b)

Kempthorne y Nordskog (1959), según se discutió anteriormente, propusieron un índice que establece restricciones basadas en un nivel de mejora predeterminada, y desarrollaron la teoría de los índices de selección restrictivos. En un índice, si uno o más caracteres se mantienen sin cambios (sin ganancia genética), mientras que el resto de caracteres muestran ganancia genética (positiva o negativa), entonces se habla de un índice restrictivo (Cunningham *et al.*, 1970; Bulmer, 1980; Cerón-Rojas *et al.*, 2008b). Las propiedades del índice KN son similares al de Smith, requieren de valoraciones económicas para establecer ponderaciones a cada carácter. Las desventajas de los índices de Smith y KN son las mismas: requiere de grandes cantidades de información, las ponderaciones económicas son difíciles de establecer según se aprecia en el trabajo de Jackson *et al.*, (2015), los errores de muestreo pueden ser muy grandes, y las propiedades de muestreo y las respuestas a la selección son desconocidas (Cerón-Rojas *et al.*, 2008b).

Cerón-Rojas *et al.* (2006) propusieron una opción a un sistema de cálculo de índices de selección basado en un ACP, pero que además incluye la posibilidad de hacer restricciones. Como se presentó en el acápite anterior, ESIM es también basado en un ACP; sin embargo tiene dos limitaciones: asume que el vector de coeficientes del índice (**b**) es igual a la matriz de varianzas/covarianzas genotípicas (**G**) multiplicadas por los ponderadores económicos desconocidos (**v**), esto es (**Gv=p**), y no permite realizar restricciones sobre algún número de caracteres (Cerón-Rojas *et al.*, 2008b). Una propuesta diferente para un índice restrictivo fundamentado, también en un ACP como ESIM, fue presentada por Cerón-Rojas *et al.*, (2008b), el índice se nombró RESIM. De manera similar a KN, RESIM asigna ponderadores a los caracteres, basado en las matrices de covarianzas entre caracteres. En RESIM se asume que el vector de ponderadores económicos comprende cantidades variables que pueden ser estimadas (Cerón-Rojas *et al.*, 2008b).

Siguiendo con la propuesta del índice KN, los m caracteres pueden no utilizarse todos, de tal forma que se restrinjan algunos de estos atributos (r); así, solamente $m-r$ atributos serán mejorados y se dejan r atributos sin cambios. La matriz $m \times r$ (\mathbf{W}) de ceros y unos está estructurada de tal manera que la varianza de dicha matriz ($\mathbf{b}'\mathbf{G}\mathbf{W}$) es igual a cero (en donde los unos son usados para las características que se mantienen sin cambios o fijas). Si se define a $\mathbf{C} = \mathbf{G}\mathbf{W}$, es posible maximizar $\mathbf{a}'\mathbf{G}\mathbf{b}$, “ \mathbf{a} ” fue definida en el acápite sobre el índice KN. La idea general de RESIM es la maximización, por medio de un ACP, de la correlación entre el índice (\mathbf{Y}) y el valor genotípico (\mathbf{Z}) ($\rho_{\mathbf{YZ}}$) bajo las siguientes condiciones: $\beta'\mathbf{S}\beta = 1$ y $\beta'\mathbf{C} = 0$ (Kempthorne y Nordskog, 1959) y, además, por $\mathbf{a}'\mathbf{G}\mathbf{a} = 1$ (Cerón-Rojas *et al.*, 2008b); entonces el valor de la covarianza $\mathbf{v}'\mathbf{G}\mathbf{b}$ puede ser maximizado bajo dichos condicionamientos (Cerón-Rojas *et al.*, 2008b). Así pues es necesario maximizar Φ con respecto a \mathbf{b} , \mathbf{v} , $\mathbf{u}' = [u_1, \dots, u_r]$, 0.5μ y 0.5ω .

$$\Phi = \mathbf{a}'\mathbf{G}\mathbf{b} - \mathbf{u}'\mathbf{C}'\mathbf{b} - 0.5\mu(\mathbf{b}'\mathbf{P}\mathbf{b} - 1) - 0.5\omega(\mathbf{a}'\mathbf{G}\mathbf{a} - 1) \quad 1.26$$

Donde \mathbf{b}' es el vector de coeficientes de RESIM, \mathbf{a}' es el vector de ponderaciones económicas, \mathbf{u}' es el vector de variables fijas o sin cambios y 0.5μ y 0.5ω son los multiplicadores de Lagrange (Cerón-Rojas, *et al.*, 2008b).

Los valores que maximizan $\mathbf{a}'\mathbf{G}\mathbf{b}$ bajo las tres condicionantes antes expuestas son: **a)** el primer (más elevado) *eigenvalue* y **b)** el vector que permite la construcción de $\mathbf{Y} = \mathbf{b}'\mathbf{p}$; en RESIM es el primer *eigenvalue* y el primer *eigenvector* (\mathbf{b}) de la matriz \mathbf{Q} ($\mathbf{Q} = [\mathbf{I} - \mathbf{P}^{-1}\mathbf{C}(\mathbf{C}'\mathbf{P}^{-1}\mathbf{C})^{-1}\mathbf{C}']\mathbf{P}^{-1}\mathbf{G}$). Los valores de ponderación (\mathbf{a}) pueden ser definidos como $\mathbf{a} = \mathbf{G}^{-1}(\mu\mathbf{P}\mathbf{b} + \mathbf{C}\mathbf{u})$. En términos de un ACP, con respecto a los coeficientes \mathbf{b} , se tiene la siguiente ecuación:

$$(\mathbf{Q} - \mu^2\mathbf{I})\mathbf{b} = 0, \quad 1.27$$

La Ecuación 1.27 es correspondiente con las ecuaciones 1.20 y 1.24, y sus componentes fueron definidos con respecto a la Ecuación 1.20.

Para la estimación de los parámetros de RESIM se debe tener en cuenta: **a)** Sea $\mathbf{Y}_{\text{RESIM}} = \mathbf{b}'_{\text{RESIM}}\mathbf{p}$ el índice de selección restringido (ISR), construido en el contexto de RESIM, donde $\mathbf{b}_{\text{RESIM}}$ es el primer *eigenvector* del primer (más elevado) *eigenvalue* de la matriz \mathbf{Q} . En

general, los estimadores $\hat{\mu}^2$ y \hat{b}_{RESIM} son los estimadores de máxima verosimilitud del *eigenvalue* y *eigenvector* de la matriz $\mathbf{Q}\mathbf{Q}'$, respectivamente. Lo anterior permite estimar Y_{RESIM} como $\hat{Y}_{RESIM} = \hat{b}'_{RESIM}p$. **b)** La respuesta a la selección (\mathbf{R}) puede ser enunciada como $R_{RESIM} = k\mu$, que corresponde al producto del diferencial de selección (k en unidades estándar) por la desviación estándar del valor del *eigenvalue* más grande.

El desarrollo teórico en detalle de RESIM y ejemplos de su utilización se encuentran en Cerón-Rojas *et al.* (2008b).

1.5.6 El mejor estimador lineal insesgado (Best linear unbiased predictor), BLUP

1.5.6.1 Modelos Lineales Mixtos, BLUE y BLUP

En mejoramiento genético de plantas es común la medición de muchas variables asociadas a caracteres de importancia económica, en ensayos de evaluación, con un modelo lineal para un diseño experimental específico. Estos modelos pueden ser asumidos como modelos lineales mixtos representados en notación matricial de la siguiente forma (Henderson *et al.*, 1959; Robinson, 1991; Piepho *et al.*, 2008; Thompson, 2008):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad 1.28$$

donde \mathbf{y} es el vector de observaciones para una variable en particular (o conjunto de variables), $\boldsymbol{\beta}$ y \mathbf{u} son los vectores para los efectos fijos y aleatorios respectivamente, en tanto que \mathbf{X} y \mathbf{Z} son las matrices incidentes conocidas del diseño y, “ \mathbf{e} ” es el vector aleatorio de residuos (error). El modelo se denomina mixto porque incluye efectos fijos (el vector de intercepto o medias, y en este ejemplo el vector $\boldsymbol{\beta}$), y además incluye efectos aleatorios (el vector de residuos \mathbf{e} , y en este caso al vector \mathbf{u}).

Se asume que los efectos aleatorios están distribuidos como $\mathbf{u} \sim \mathbf{MVN}(\mathbf{0}, \mathbf{G})$ y $\mathbf{e} \sim \mathbf{MVN}(\mathbf{0}, \mathbf{R})$, donde, generalizando, $\mathbf{MVN}(\boldsymbol{\mu}, \mathbf{V})$ denota la distribución normal multivariada (*Multivariate Normal Distribution*) con media $\boldsymbol{\mu}$ y la matriz incidente \mathbf{V} de varianzas/covarianzas. Los efectos fijos en el modelo de la Ecuación 1.28 pueden ser estimados por medio de *Best Linear Unbiased Estimation* (**BLUE**), en tanto que los efectos aleatorios pueden ser estimados por *Best Linear Unbiased Predictor* (**BLUP**) (Piepho *et al.*,

2008). Las varianzas en el modelo en la Ecuación 1.28 se denotan como sigue (Thompson, 2008):

$$\text{Var}(\mathbf{y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R} = \mathbf{V};$$

$$\text{Var}(\mathbf{u}) = \mathbf{G};$$

$$\text{Var}(\mathbf{e}) = \mathbf{R}.$$

donde, las matrices \mathbf{V} , \mathbf{G} y \mathbf{R} representan las matrices de varianzas de \mathbf{y} , \mathbf{u} y \mathbf{e} , respectivamente.

En otros términos (Robinson, 1991), puede indicarse que las esperanzas matemáticas de las medias de los vectores de los efectos aleatorios $E(\mathbf{u})$ y $E(\mathbf{e})$ son iguales a cero, y las varianzas de dichos vectores son (comparar con $MVN(\boldsymbol{\mu}, \mathbf{V})$) :

$$\text{Var} \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \sigma^2 \quad 1.29$$

La estimación de $\boldsymbol{\beta}$ y \mathbf{u} por medio de **BLUE** y **BLUP**, respectivamente, se hace mediante los estimadores $\hat{\boldsymbol{\beta}}$ y $\hat{\mathbf{u}}$, también llamados las “soluciones de las ecuaciones de los modelos lineales mixtos”.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \quad 1.30$$

$$\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad 1.31$$

Un problema con la solución de las ecuaciones 1.30 y 1.31 es la inversión de la matriz de varianzas \mathbf{V} de \mathbf{y} a \mathbf{V}^{-1} . Henderson *et al.* (1959) mostraron que la inversión de la matriz \mathbf{V} , que es la matriz de varianzas de $n \times n$ (n = al número de observaciones i), puede evitarse y, además, se mantiene que $\hat{\boldsymbol{\beta}}$ satisface a:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad 1.32$$

En la expresión 1.32 se presentan los términos, tanto de los efectos fijos ($\boldsymbol{\beta}$) como de los efectos aleatorios (\mathbf{u}). La expresión 1.32 en realidad son dos ecuaciones que son conocidas como Ecuaciones de un Modelo Mixto (MME, por sus siglas en inglés). Como se apuntó arriba, estas ecuaciones pueden ser resueltas mediante el cómputo de **BLUE** y **BLUP**. Para

factores como el componente genético, si éste es considerado como un efecto fijo, Henderson *et al.* (1959) indican que la estimación de dicho efecto, por medio de la Máxima Verisimilitud (*Maximum Likelihood*, en inglés), conduce a estimaciones sesgadas. Las estimaciones de **BLUP** generalmente difieren de las estimaciones por mínimos cuadrados generalizados en el caso de que **u** fuera considerado un efecto fijo (Robinson, 1991). Por lo anterior, según el mismo autor, el uso de **BLUP** se presenta como una opción para la solución de este problema, siempre que el efecto sea reconsiderado como aleatorio. Una discusión amplia sobre el origen de este sesgo, en el uso de los mínimos cuadrados en la Máxima Verosimilitud, es discutida por Kempthorne y Krosigk en Henderson *et al.* (1959).

En general, el interés principal en el análisis de experimentos de materiales genéticos es la estimación del efecto de los tratamientos, tomando en consideración la matriz de varianzas correspondiente. Algunas aplicaciones genéticas, sobre un modelo lineal mixto, están relacionadas con el interés en la estimación de la matriz de varianzas/covarianzas **G** (que, en general, son las varianzas/covarianzas del valor genético de las observaciones); o bien el ajuste de los datos al efecto fijo **β** . Sin embargo, una aplicación en particular de mucho interés es también la *predicción* de los efectos aleatorios de **u** (Thompson, 2008).

Una propiedad importante de **BLUP** es la *aproximación* del **BLUP** hacia la media de la variable (*y*) de la cual se deriva. Lo anterior anticipa una regresión de la progenie con respecto a la media observada de *y* (Thompson, 2008; Piepho *et al.*, 2008); sin embargo, la regresión del mérito genético de un material con respecto a la media de *y* (\bar{y}) dependerá de la cantidad de información que se tenga de dicho material genético; la aproximación de los componentes **BLUP**, con respecto a la media, es menos esparcida de lo que sería si los componentes de **u** hubiesen sido considerados como efectos fijos (Robinson, 1991). En todo caso esta propiedad se contempla como una ventaja de **BLUP**. **BLUP**, además, se considera insesgado por cuanto $\hat{u} - u = 0$ (Robinson, 1991; Thompson, 2008).

Un aspecto importante sobre los modelos mixtos y sus estimadores **BLUE** y **BLUP** es la dificultad de distinguir entre efectos fijos y efectos aleatorios. En general, la diferenciación o la decisión sobre considerar o no como efecto aleatorio o fijo a un factor, puede resultar difícil. Robinson (1991) indica que un efecto debe considerarse aleatorio si este proviene de una población con una distribución de probabilidades, de lo contrario debe considerarse un efecto fijo.

Por otro lado, Robinson (1991) indica que para variables de clase (bloques, sitios o años), si los efectos de estos no son confundidos (por manejo completamente diferencial por ejemplo), entonces estos deben considerarse como efectos aleatorios. Pero si los efectos no pueden ser deslindados según los niveles para cada variable de clase, entonces, estos deben considerarse como efectos fijos. Esta última situación es más común en experimentación con animales que con plantas. Por lo que, para mejoramiento genético en plantas, prácticamente todos los factores pueden llegar a ser considerados como aleatorios, ya que en general se guarda un buen nivel de conocimiento de dichos efectos y, por tanto, las lecturas provienen de poblaciones con distribución de probabilidad.

La distinción entre efectos fijos y aleatorios ha sido de amplio debate, y abordado por varios autores (Nelder, 1977; Robinson, 1991; Thompson, 2008; Piepho *et al.*, 2008). En varios casos ha sido un tema de controversia. En los últimos años, con el advenimiento de nuevos sistemas de análisis por computación, este tema se ha superado en parte, y las aplicaciones de los modelos lineales mixtos en plantas es de mayor uso, en particular en el mejoramiento genético (Simeão *et al.*, 2002; Moreira *et al.*, 2005; Piepho y Möhring, 2006; Da Silva *et al.*, 2016).

Robinson (1991) dice, con respecto a la distinción entre efectos aleatorios, y citando al Dr. Jhon W. Tukey en la discusión de la lectura de Nelder (1977), "...el enfoque debe ser fundamentado en preguntas y no en modelos...". La elección de si un factor debe ser tratado como fijo o aleatorio debe, entonces, fundamentarse en las preguntas que estamos tratando de responder. Esto debería incluir las hipótesis que se traten de probar.

Detalles sobre álgebra y la teoría fundamental sobre los modelos lineales mixtos se presentan en Henderson *et al.*, 1959; Robinson, (1991); Thompson (2008); Bates, (2013).

1.5.6.2 BLUP multi-carácter o MT-BLUP (*Multi Trait BLUP*)

La discusión anterior y la relación sobre el sistema de ecuaciones para modelos lineales mixtos discutidos en el epígrafe anterior corresponden con el análisis o estudio de un solo carácter o variable. Sin embargo, este puede ser considerado en realidad un caso de **BLUP** multi-carácter (**MT-BLUP**, *Multi Trait BLUP*). Las ecuaciones del modelo lineal mixto, para un solo carácter, fueron presentadas en la expresión 1.32. Un modelo multi-carácter puede considerarse como una extensión de un modelo uni-carácter (ya discutido); o bien este último

puede considerarse un caso del modelo multi-carácter. En general, el análisis **MT-BLUP** produce resultados de los valores genéticos para cada uno de los caracteres en estudio; sin embargo, los resultados son, generalmente, diferentes a los que se podrían obtenerse si los análisis se realizaran individualmente para cada carácter (Van der Werf, 2002). Lo anterior se debe a que en el **MT-BLUP** cada característica es usada para ayudar a dar información acerca de todos los otros caracteres, de una forma similar a la que se utiliza en el cómputo de un índice de selección (Van der Werf, 2002; Colleau *et al.*, SF).

Las ventajas de un **MT-BLUP** se deben, principalmente, a una mayor precisión conforme se utiliza información de caracteres correlacionados, y a un menor sesgo conforme el análisis considere caracteres que son medidos en rondas secuenciales de selección: solamente los mejores genotipos son evaluados (Van der Werf, 2002). Cuando se evalúan secuencialmente, solamente las mejores progenies son seleccionadas, lo que genera un sesgo, ya que se les puede brindar ventajas a los malos progenitores. **MT-BLUP** corrige este sesgo.

Para la evaluación genética de un material puede emplearse la información contenida en todas las variables en estudio. Con el uso de la información de caracteres correlacionados (en caña de azúcar pueden ser las variables asociadas al contenido de azúcar), la precisión de la estimación de los valores genéticos (*breeding values*) se incrementa (Van der Werf, 2002).

El **MT-BLUP** fue originalmente propuesto por Henderson y Quaas (1976). El procedimiento se resume así (Henderson y Quaas, 1976; Van der Werf, 2002): para un análisis uni-carácter se retoma la Ecuación 1.28, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$, donde \mathbf{y} es el vector de observaciones para “una” variable o carácter.

Con más caracteres se podría partir el vector de observaciones “ \mathbf{y} ” en partes proporcionales para cada uno de los caracteres. Lo mismo puede hacerse con los efectos ambientales asociados (efectos fijos, que ahora es $\boldsymbol{\beta}$); en tanto que el vector para los valores genéticos (efectos aleatorios, que ahora es \mathbf{u}), también puede ser partido para los diferentes caracteres, de tal forma que cada tratamiento o material genético tiene un valor genético para cada uno de los caracteres. Por simplicidad, considerando solamente dos caracteres, el vector \mathbf{y}_1 representa la \mathbf{n}_1 observación para el carácter “1”, en tanto que \mathbf{y}_2 representa la \mathbf{n}_2 observación para el carácter “2”. Por tanto, de forma similar a la Ecuación 1.28, para cada carácter se puede escribir un modelo mixto de la siguiente forma:

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{Z}_i \mathbf{u}_i + \mathbf{e}_i \quad 1.33$$

En la Ecuación 1.33 hay implicados p efectos fijos asociados a con el carácter i , de tal forma que \mathbf{X}_i es una matriz $n_i \times p_i$ y $\boldsymbol{\beta}_i$ es una matriz $p_i \times 1$ (vector columna). \mathbf{X}_i y \mathbf{Z}_i son las matrices incidentes para los efectos fijos y aleatorios para la característica “ i ”, respectivamente. El modelo multi-carácter puede ser representado como sigue:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 \\ 0 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & 0 \\ 0 & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \quad 1.34$$

Un aspecto importante es que no necesariamente todos los tratamientos (genotipos) deben contar con observaciones para los dos caracteres. Sin embargo, al final, todos los tratamientos se representan con un valor genético para cada característica en el análisis, sin importar si contaban o no con información acerca de ese carácter. Esta es justamente una de las ventajas de un ajuste de un modelo mixto multi-caracteres (Henderson y Quaas, 1976). Por lo anterior, los vectores \mathbf{y}_1 y \mathbf{y}_2 no necesariamente son del mismo tamaño, al igual que sus errores asociados (\mathbf{e}_i). De forma análoga que para un modelo lineal uni-carácter, es necesario especificar las matrices de varianzas/covarianzas R (Ecuación 1.36) y G (Ecuación 1.35), asociadas al vector de los residuos (\mathbf{e}_i) y el vector de los efectos aleatorios (en este caso \mathbf{u}_i), respectivamente (Henderson y Quaas, 1976).

$$Var(u) = G = \begin{bmatrix} g_{11}A & g_{12}A & \dots & g_{1q}A \\ g_{12}A & g_{22}A & \dots & g_{2q}A \\ \vdots & \vdots & \ddots & \vdots \\ g_{q1}A & g_{q2}A & \dots & g_{qq}A \end{bmatrix} = G_0 * A \quad 1.35$$

donde, A es la matriz numerador $n \times n$ de relaciones entre tratamientos (genotipos), “*” hace referencia a la dirección correcta de operación de la producción, en tanto que g_{ij} es el ij -ésimo elemento de G_0 (Henderson y Quaas, 1976).

$$Var(e) = R = \begin{bmatrix} r_{11}I & r_{12}I & \dots & r_{1q}I \\ r_{12}I & r_{22}I & \dots & r_{2q}I \\ \vdots & \vdots & \ddots & \vdots \\ r_{q1}I & r_{q2}I & \dots & r_{qq}I \end{bmatrix} = R_0 * I \quad 1.36$$

Donde, r_{ij} es el ij -ésimo elemento de R_0 , I es una matriz identidad $n \times n$.

Las ecuaciones multi-caracteres de modelos lineales mixtos crecen abundantemente según el número de caracteres a analizar. Es por esta razón que en 1976, Henderson y Quaas (1976) indicaban que era preciso establecer un balance entre la necesidad de obtener las ventajas de la resolución de modelos multi-carácter y los grandes problemas computacionales implicados. En la actualidad, los sistemas computacionales automatizados han resuelto en buena medida este problema, y su aplicación en la mejora genética en modelos multi-carácter es posible. En la Ecuación 1.37 se presenta un ejemplo de Ecuaciones de Modelos Lineales Mixtos para dos caracteres únicamente. Se aprecia en esta expresión la complejidad que se alcanza, y es solo el ejemplo para dos caracteres. La forma general de estas ecuaciones es presentada por Henderson y Quaas (1976).

$$\begin{bmatrix} \mathbf{X}_1' \mathbf{r}^{11} \mathbf{X}_1 & \mathbf{X}_1' \mathbf{r}^{12} \mathbf{X}_2 & \mathbf{X}_1' \mathbf{r}^{11} \mathbf{Z}_2 & \mathbf{X}_1' \mathbf{r}^{12} \mathbf{Z}_2 \\ \mathbf{X}_2' \mathbf{r}^{21} \mathbf{X}_1 & \mathbf{X}_2' \mathbf{r}^{22} \mathbf{X}_2 & \mathbf{X}_2' \mathbf{r}^{21} \mathbf{Z}_1 & \mathbf{X}_2' \mathbf{r}^{22} \mathbf{Z}_2 \\ \mathbf{Z}_1' \mathbf{r}^{11} \mathbf{X}_1 & \mathbf{Z}_1' \mathbf{r}^{12} \mathbf{X}_2 & \mathbf{Z}_1' \mathbf{r}^{11} \mathbf{Z}_1 + \mathbf{g}^{11} \mathbf{A}^{-1} & \mathbf{Z}_1' \mathbf{r}^{12} \mathbf{Z}_2 + \mathbf{g}^{12} \mathbf{A}^{-1} \\ \mathbf{Z}_2' \mathbf{r}^{21} \mathbf{X}_1 & \mathbf{Z}_2' \mathbf{r}^{22} \mathbf{X}_2 & \mathbf{Z}_2' \mathbf{r}^{21} \mathbf{Z}_1 + \mathbf{g}^{21} \mathbf{A}^{-1} & \mathbf{Z}_2' \mathbf{r}^{22} \mathbf{Z}_2 + \mathbf{g}^{22} \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1' (\mathbf{r}^{11} \mathbf{y}_1 + \mathbf{r}^{12} \mathbf{y}_2) \\ \mathbf{X}_2' (\mathbf{r}^{21} \mathbf{y}_1 + \mathbf{r}^{22} \mathbf{y}_2) \\ \mathbf{Z}_1' (\mathbf{r}^{11} \mathbf{y}_1 + \mathbf{r}^{12} \mathbf{y}_2) \\ \mathbf{Z}_2' (\mathbf{r}^{21} \mathbf{y}_1 + \mathbf{r}^{22} \mathbf{y}_2) \end{bmatrix} \quad 1.37$$

Con respecto a las aplicaciones de los modelos mixtos en un programa de mejoramiento genético, y en el desarrollo de las nuevas variedades en plantas (como el que se conduce en CENGICANÑA) en estados iniciales de selección, los efectos relacionados con el componente genético suelen ser considerados como un efecto fijo y, por tanto, parte de β en un modelo lineal mixto. Sin embargo, cuando el factor *genotipos* se considera como aleatorio, los efectos genotípicos pasan a formar parte de \mathbf{u} , y entonces son estimados por **BLUP** (Piepho *et al.*, 2008), con lo cual se obtiene un valor insesgado del valor genético de los materiales en estudio; es decir, una especie de índice de desempeño.

Los efectos genéticos son aquellos que están relacionados con la descomposición del valor genético en componentes de efectos con interpretación genética en términos de caracteres de variación cuantitativa (Bos y Caligari, 2007; Piepho *et al.*, 2008). Estas descomposiciones en términos fenotípicos se presentaron en la Tabla 1.3. Para el caso de la caña de azúcar, que es de reproducción vegetativa, la descomposición del fenotipo se restringe al valor genético general más el efecto del ambiente ($p = g + e$) (Tabla 1.3) (Bos y Caligari, 2007). Para el caso de ensayos de cruza dialélicas la descomposición puede realizarse en términos de la habilidad combinatoria general y específica (Piepho *et al.*, 2008). Por tanto, la

aspiración principal en un programa de mejora genética es la predicción del valor genotípico de un genotipo candidato (Piepho *et al.*, 2008).

Para los efectos de este trabajo, los valores de **BLUP** se consideraron como los estimadores del valor genético de los genotipos en evaluación y se utilizaron como índice de desempeño en términos genéticos con respecto a la variable Toneladas de Azúcar por Hectárea, TAH.

1.5.7 Toneladas de Azúcar por Hectárea, TAH

El TAH no es un índice *per se*. No es, necesariamente, una función lineal de algún conjunto de variables o factores. Más bien es el resultado de operaciones aritméticas de varias mediciones hechas en el campo (Rosales-Longo, 2016b). En la sección de Materiales y Métodos se presenta el procedimiento utilizado en este trabajo para la obtención de este valor. El TAH es producto de la acumulación de biomasa y la concentración de azúcar en los culmos de la caña de azúcar (Ming *et al.*, 2006). Su medida está dada por la producción de biomasa (caña) en $t\ ha^{-1}$ y la concentración relativa (porcentaje) de sacarosa en los culmos (Orozco *et al.*, 2012). La concentración relativa también puede estar dada por la cantidad de kilos estimados de azúcar por tonelada de caña (Pérez *et al.*, 2016).

El TAH es una medida de amplio uso en todas las agroindustrias azucareras alrededor del mundo y sirve para estimar la productividad general del cultivo de la caña de azúcar. En mejoramiento genético de la caña de azúcar, es la medida final que permite establecer el valor genético de los materiales en evaluación y es, además, una importante medida de discriminación entre genotipos (Orozco *et al.*, 2012). Si bien otras variables como la oquedad de los tallos, el corcho en los tallos, la fibra y el nivel de resistencia a enfermedades son de importancia económica, el TAH es de alta ponderación entre los fitomejoradores en el momento de tomar decisiones finales de selección (Berding *et al.*, 2004; Ming *et al.*, 2006; Orozco *et al.*, 2012).

La concentración de sacarosa puede ser medida en diferentes etapas de la producción de caña de azúcar. Pueden tomarse mediciones previas a la cosecha, por medio de lecturas refractométricas *in situ*; o bien puede tomarse una muestra antes de la cosecha y posteriormente evaluar, en condiciones de laboratorio, la concentración y pureza de la sacarosa en los jugos de la muestra (Rosales-Longo, 2016b). La forma del muestreo puede

variar. En las operaciones industriales la muestra de caña para la estimación de la concentración de azúcar se realiza en el patio de descarga, mediante el uso de una sonda (*core sampler*), la cual se introduce en el contenedor para extraer una muestra. La mayoría de las fábricas utilizan un sistema similar para establecer el potencial de azúcar en la caña que ingresa a las instalaciones fabriles (Birkett, 1998). En este sistema, las mediciones tradicionales son: una prueba de sedimentos, °Brix del jugo (refractometría), Pol en jugo, y el porcentaje de humedad residual (Birkett, 1998).

Todas las medidas anteriores son utilizadas para el cálculo final de variables de utilización práctica como el Pol % Caña. En este trabajo, la variable utilizada para el cálculo de TAH es Pol % Caña. El Pol % Caña es una medida derivada de la medición de la concentración de azúcar en un aparato llamado polarímetro, el cual mide la rotación de un haz de luz polarizada proyectada en una columna conteniendo la muestra de jugo (Engelke, 2002). Se emplean también, para la medición, estándares y tablas (Engelke, 2002). El producto final de la lectura es la medida llamada Pol % Jugo, la cual se asocia con el contenido de materia seca para finalmente obtener Pol % Caña. La medida de la concentración de sacarosa en Pol % Caña es multiplicada por la medida en fresco de biomasa (puede ser en toneladas), para obtener finalmente la cantidad de azúcar estimada en un área en particular, hectáreas por ejemplo, es decir, el TAH.

La caracterización final de las variedades de caña de azúcar incluye al menos las siguientes variables: Azúcar por unidad de área, Biomasa por unidad de área, y contenido relativo de sacarosa. Las variables pueden ser diferentes según el programa de mejora genética; por ejemplo, en Guatemala se miden: Toneladas Métricas de Caña por Hectárea (TCH), Toneladas de Azúcar por Hectárea (TAH) y Pol % Caña (Orozco *et al.*, 2012). En Estados Unidos las dimensionales son: Toneladas Inglesas de Caña por Acre (TCA), Libras de Azúcar por Acre y Libras de Azúcar por Tonelada Inglesa de Caña (LSU, 2014).

El rendimiento ha sido un carácter que los programas de mejoramiento buscan desarrollar positivamente (Allard, 1980; Fehr, 1987; Ming *et al.*, 2006; James, 2004). En este sentido, una de las medidas de rendimiento en caña de azúcar es el TAH. En general, un progreso en los valores del rendimiento es una medida apropiada para establecer el nivel de ganancias genéticas en un programa de mejora genética. En el presente estudio, el TAH fue utilizado más como un indicador del desempeño general de los materiales genéticos, que como un

componente de las funciones lineales estudiadas. No se empleó en el desarrollo de las funciones lineales de los índices en estudio ya que, al ser una variable derivada de varias de las 14 variables utilizadas en este estudio, genera colinealidad y provoca la posibilidad de llegar a conclusiones erróneas debido al sesgo que supone la redundancia en el efecto de algunas de las variables utilizadas para la generación de los índices de selección.

1.6 PROBLEMA ESPECÍFICO DEL ESTADO III DEL PROGRAMA DE VARIEDADES DE CENGICAÑA

La información sobre los resultados de investigación de selección en el Estado III del Programa de Variedades de CENGICAÑA no se explota en todo su potencial. Estos resultados se utilizan de forma discreta, con análisis estadísticos de frecuencias y de variación (Montgomery, 1991), lo que puede limitar las actividades de selección. Además, puede conllevar algún grado de riesgo en la selección del material genético, dando como resultado la selección de genotipos de mediano potencial de desempeño, o bien la omisión de genotipos con un alto potencial de rendimiento en biomasa y azúcar. En consecuencia, la selección de los materiales genéticos tiende a ser un poco menos precisa con el consecuente consumo de tiempo, esfuerzo y otros recursos, lo que puede limitar la posibilidad de obtener los materiales genéticos requeridos, con las mejores características de adaptación a los sistemas de producción de caña de azúcar en Guatemala.

1.7 DISCUSIÓN Y JUSTIFICACIÓN DEL PROYECTO RESPECTO AL PROBLEMA ESPECÍFICO

En caña de azúcar, como en otros cultivos, el conocimiento del cultivo y el perfecto conocimiento del tipo buscado es la clave de una buena selección. Sin embargo, un tipo debe reunir una gran cantidad de atributos que pueden comprender distintas variables de valoración y de clasificación (Tabla 1.4).

Tabla 1.4. Algunas características que debe reunir una variedad de caña de azúcar.

Característica
1. Alta producción de biomasa (población, altura, diámetro del tallo, peso específico).
2. Resistencia a enfermedades (escaldadura, carbón, royas, raya roja, mosaico, amarillamiento, otras).
3. Resistencia a herbicidas.
4. Poca o ninguna flor.
5. Adecuada producción de fibra.
6. Alta eficiencia en la asimilación de CO ₂ .
7. Alta calidad de los jugos.
8. Buen rebrote.
9. Alta concentración de sacarosa.
10. Resistencia a plagas, (barrenador, chinche salivosa, chinche de encaje, rata, coludo, cornudo, otras).
11. Resistencia a maduradores.
12. Poco o ningún corcho.
13. Baja producción de material extraño (trash).
14. Bajo nivel de producción de almidones.
15. Apropiada para la cosecha mecanizada.
16. Buen cierre entre surcos para el control de malezas.
17. Otras.

Si se considera que en los diferentes estados de selección se evalúan una gran cantidad de genotipos, y a cada uno hay que observarle las variables pertinentes, el grado de precisión de la selección se ve menoscabado por la dificultad de realizar una valoración correcta de cada uno de los genotipos en evaluación. Por tanto, si fuera posible reducir o compilar todos los valores de cada una de las variables para cada uno de los genotipos, cada uno ponderado según su valor de importancia, entonces sería posible evaluar y clasificar a los genotipos en estudio de una manera mucho más eficiente y rápida. Esto debería conducir a un proceso de selección mucho más eficiente, con resultados más precisos y de mejor calidad.

En general se busca seleccionar aquellos genotipos con altos valores genotípicos (g) para las características estudiadas (las cuales no son directamente observadas), y que están relacionadas con los valores fenotípicos observables o medibles (p). Se asume que los valores genotípicos no observables y sus valores fenotípicos observables tienen una probabilidad conjunta de distribución, de tal manera que si se realiza una regresión entre los valores genotípicos y fenotípicos, los valores de dicha regresión pueden considerarse un buen estimador de g y, por tanto, los genotipos con los valores más altos de dicha regresión pueden ser sujetos de selección (Cerón-Rojas *et al.*, 2006). Lo anterior es el principio de un índice de desempeño o de selección. El cálculo correcto de una función de regresión brindará un buen

estimador de los valores genéticos de los genotipos en evaluación y, por tanto, un valor insesgado que permita una selección eficiente y objetiva.

Sin embargo, determinar los estimadores de los valores genéticos de los materiales en evaluación es complicado cuando existe un alto nivel de interacción entre el genotipo y el ambiente, tal como ocurre con la caña de azúcar. Por lo tanto, es necesario realizar una evaluación de procedimientos que permitan reducir a su mínima expresión esa interacción, de tal manera que sea posible el cálculo de los estimadores genéticos (índices de selección) de alta utilidad práctica.

Hoy en día los estimadores genéticos en el Programa de Variedades de CENGICAÑA no están desarrollados o son muy limitados en su uso, en parte porque los estudios básicos de desarrollo de los mismos no habían sido realizados. El establecimiento de estos estudios y los resultados de los mismos permitirán al Programa de Variedades mejorar sustancialmente su programa de selección y, por consiguiente, el esquema general de trabajo se verá mejorado y, en consecuencia, los resultados del mismo tendrán que ser de una calidad cada vez más optimizada.







2. OBJETIVOS E HIPÓTESIS



2. OBJETIVOS E HIPÓTESIS

2.1 OBJETIVOS

Este trabajo buscó, en general, aplicar procedimientos avanzados desde el punto de vista del mejoramiento genético, a fin de perfeccionar los procesos de selección de los mejores genotipos de caña de azúcar. Para poder llevar a buen término lo anterior, fue necesaria la inspección fenotípica detallada de los materiales, utilizando tecnologías y procedimientos estadísticos avanzados para llevar a un nuevo nivel el trabajo del mejoramiento genético.

Además, con lo anterior, el Programa de Variedades permitirá a la agroindustria azucarera guatemalteca contar con variedades de superior desempeño en un plazo razonable de tiempo, reduciendo el riesgo de selecciones erróneas y, por tanto, será posible también contar con un mayor número de materiales genéticos en los estados avanzados de selección. Con lo anterior se aumenta la probabilidad de liberar variedades de caña de azúcar con adaptación específica, pero también se incrementa la probabilidad de encontrar genotipos de adaptación general.

Adicionalmente, este trabajo tuvo por objetivo identificar fenotípicamente (p) aquellas variables que estaban más íntimamente relacionadas con los valores genéticos (g) de los genotipos; y, por tanto, establecer la utilidad real de todas las mediciones que actualmente se hacen, de tal manera que, si se establece con buena precisión, también será posible ahorrar tiempo en el campo, en el laboratorio y en el gabinete.

2.1.1 Objetivo General

Con este trabajo se procuró el desarrollo y adaptación de las metodologías que permitieran un trabajo de mejoramiento más eficiente en el Programa de Variedades de CENGICAÑA. En consecuencia, los resultados de este trabajo, además de aportar un cúmulo de información en conocimiento específico, se pretendió que fueran de aplicación práctica y rutinaria en el Programa de Variedades de CENGICAÑA.

Por otro lado, se buscó que una vez establecidos de forma rutinaria estos procedimientos, fueran de constante mejora mediante la generación de más información y la adaptación de nuevos procedimientos que continuamente procuren la mejora en el programa de Mejoramiento Genético.

2.1.2 Objetivos Específicos

- 2.1.2.1 Evaluar nueve metodologías para el cómputo de índices de desempeño o selección a partir de la información fenotípica de los genotipos de caña de azúcar del proyecto del “*Estado III de Selección*” del Programa de Variedades de CENGICAÑA en Guatemala.
- 2.1.2.2 Determinar el grado de ajuste de los modelos en estudio para establecer el nivel de su utilidad práctica.
- 2.1.2.3 Identificar las principales variables (componentes de rendimiento) que determinan el desempeño de los genotipos de caña de azúcar.

2.2 HIPÓTESIS

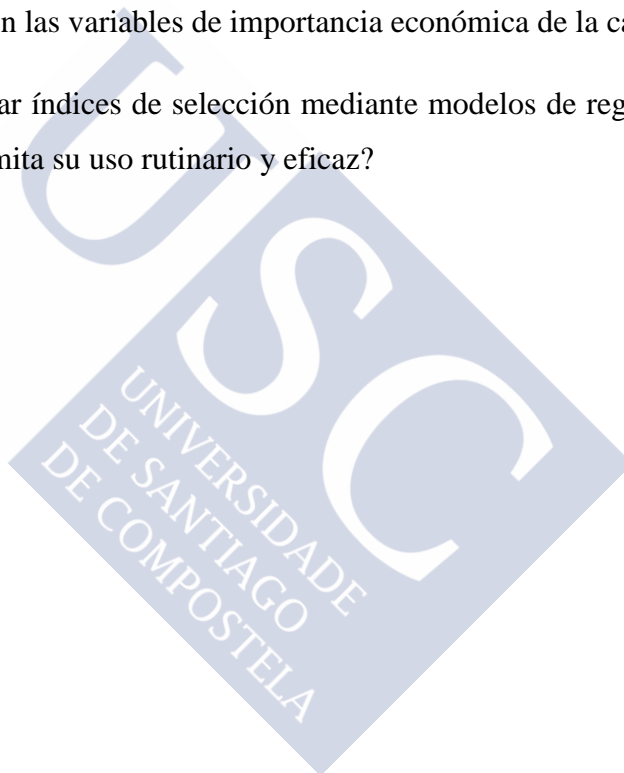
Hipótesis (nulas)

- 2.2.1 La selección eficiente de los genotipos de caña de azúcar no es función de la aplicación de índices de selección versus la selección visual individual o por competencia del fitomejorador.
- 2.2.2 Los modelos que generan los índices de selección no presentan ajustes estadísticamente aceptables y, por tanto, no son de utilidad práctica.

- 2.2.3 No es posible identificar las variables (características) que mejor se correlacionan con un índice de selección.

Con lo anterior se buscaron las respuestas a las preguntas esenciales de Tesis Doctoral, que estuvieron encaminadas al mejoramiento sustancial del Programa de Variedades de CENGICAÑA en la selección de variedades de caña de azúcar en el Estado III:

1. ¿Son capaces los índices de selección de optimizar razonablemente el proceso de mejoramiento de la caña de azúcar en Guatemala?
2. ¿Cuáles de todas las variables estudiadas en el Estado III estuvieron realmente correlacionadas con las variables de importancia económica de la caña de azúcar?
3. ¿Es posible calcular índices de selección mediante modelos de regresión con un buen ajuste, tal que permita su uso rutinario y eficaz?







3. MATERIALES Y MÉTODOS



3. MATERIALES Y MÉTODOS

3.1 EXPERIMENTOS Y CONJUNTOS DE DATOS

Se recolectaron datos agrupados en ocho conjuntos (Figura 3.1). La recolección de los datos se realizó en cuatro experimentos ubicados en cuatro localidades diferentes: **estrato litoral_flor**, **estrato litoral_no_flor**, **estrato medio_flor**, y **estrato medio_no_flor** (Figura 3.2). Estos experimentos fueron evaluados en dos épocas diferentes: Primer año, **plantía**; Segundo año, **1ª Soca**. Producto de la combinación de ubicación geográfica y años de selección, se obtuvieron ocho conjuntos de datos (Figura 3.1). Uno de los conjuntos de datos (**litoral_no_flor_1ª_soca**) se perdió por una quema delincencial, por lo que los datos no pudieron ser recabados. En consecuencia, se trabajó sobre los resultados de siete conjuntos de datos (Figura 3.2)

Los conjuntos de datos se agruparon en ocho (uno de ellos perdido por quema intencionada), porque la expresión de los genotipos fue diferencial en el espacio y el tiempo. La evaluación en detalle fue, en alta proporción, diferencial, esto es: **No todas las variedades evaluadas en el primer año de trabajo fueron evaluadas en el segundo año de trabajo y viceversa**. Lo anterior corresponde con el procedimiento rutinario de evaluación y selección para la mejora genética de la caña de azúcar en CENGICAÑA en Guatemala. En todo caso, se contó con lecturas repetidas de varios genotipos en los dos ciclos, incluidos los cultivares testigo, durante los dos ciclos de evaluación, lo cual permitió la generación de las varianzas respectivas para la realización de diversos cálculos, principalmente con respecto a los modelos que se describen más adelante. Estas varianzas también fueron de utilidad en los cálculos de heredabilidad.

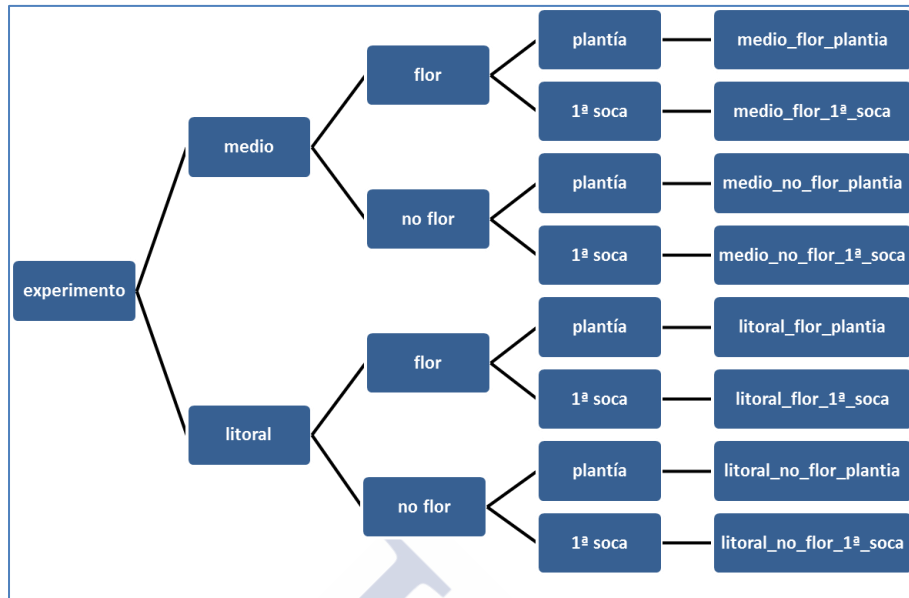


Figura 3.1. Esquema de ordenación de los conjuntos de datos sujetos a evaluación.

En resumen, la codificación y ordenamiento final de los conjuntos de datos es como sigue:

1. litoral_flor_1ª_soca
2. litoral_flor_plantía
3. litoral_no_flor_1ª_soca (perdido por un incendio delincuencial)
4. litoral_no_flor_plantía
5. medio_flor_1ª_soca
6. medio_flor_plantía
7. medio_no_flor_1ª_soca
8. medio_no_flor_plantía

Los cuatro experimentos, en términos geográficos, fueron montados y evaluados en Guatemala durante los períodos de zafra de caña de azúcar, 2014-2015 y 2015-2016, dos de ellos en el litoral (entre 10 y 12 msnm) y otros dos en el Estrato medio (entre 165 y 235 msnm) (Figura 3.2).



Figura 3.2. Ubicación geográfica de los cuatro experimentos en evaluación: a) estrato medio_flor. B) estrato litoral_flor. C) estrato medio_no_flor. D) estrato litoral_no_flor (Google Earth, SF).

El área donde se ubicó el experimento del estrato **medio_flor** (Figura 3.2a), correspondió a la Zona Agroecológica (ZAE) 18, la cual corresponde a un grupo de suelos Inceptisoles de textura arcillosa, con relieve ligeramente inclinado, y susceptibles a la erosión, el déficit de humedad va de -500 a -799.9 mm; el período seco en esa zona ocurre desde el 30 de octubre al cinco de mayo (Villatoro y Meneses, 2013). El experimento de variedades florecedoras en el litoral (Figura 3.2b) se ubicó en la zona agroecológica 12. El experimento de variedades no florecedoras del estrato medio (Figura 3.2c) correspondió con la ZAE 9, la cual corresponde a suelos Mollisoles secos, de textura gruesa, moderadamente profundos y muy permeables. El déficit de humedad es de 0 a -499.9 mm. El período seco ocurre del 10 de noviembre al 25 de abril (Villatoro y Meneses, 2013), en tanto que el experimento de variedades no florecedoras del estrato litoral (Figura 3.2.d), se ubicó en la zona agroecológica (ZAE) 12, ésta cual corresponde con suelos de texturas gruesas, moderadamente profundos y permeables, de grupo de los Mollisoles secos; el déficit hídrico en esta zona es de -900 mm (Villatoro y Meneses, 2013).

En la Tabla 3.1 se aprecian las diferencias meteorológicas entre las dos zonas donde se ubicaron los experimentos. El estrato litoral tuvo medias de temperatura, mayores que las que mostró el estrato medio, lo mismo ocurrió con la radiación solar global, la cual es mayor

también en el estrato litoral, lo mismo que la humedad relativa. Sin embargo, la precipitación promedio mensual fue mayor en el estrato medio. Por otro lado, los valores máximos y mínimos reflejan las variaciones entre localidades, particularmente en el tema de la radiación y la precipitación pluvial. El estrato medio fue más lluvioso, la media anual del estrato medio fue de 3,137 mm versus el promedio del estrato litoral 106.7 mm. En consecuencia se observa que el estrato litoral fue mucho más seco y el cultivo de los experimentos en esta zona, requirió de diez riegos, con una lámina de riego de 50 mm; tres riegos se realizaron entre los meses de julio y agosto de cada año y el resto entre los meses de diciembre y marzo. En el estrato medio se realizaron seis riegos con una lámina de 40 mm. Un riego se realizó en el mes de julio y el resto entre los meses de diciembre y marzo. El período de espaciamiento entre riegos fue de quince días.

Tabla 3.1. Condiciones climáticas, valores promedio mensuales, en las áreas de experimentación durante dos ciclos de evaluación, en el período 01/01/2014 Al 31/03/2016.

Estación meteorológica	Variable	Media	D.E.	Mín	Máx	Mediana	Suma
El Bálsamo	Temperatura °C	25.54	1.17	21.31	28.65	25.59	NA
(Estrato medio)	Radiación watts m ⁻²	17.38	3.37	1.30	25.05	17.58	148.43
	(miles)						
	Humedad relativa %	82.99	7.10	46.80	100.00	83.79	NA
	Precipitación mm	8.57	19.15	0.00	132.60	0.00	7320
San Antonio	Temperatura °C	27.15	1.11	22.91	30.20	27.22	NA
El Valle	Radiación watts m ⁻²	21.33	3.53	0.0013	2.79	2.16	182.12
(Estrato litoral)	(miles)						
	Humedad relativa %	86.46	3.83	55.91	100.00	86.24	NA
	Precipitación mm	2.92	14.67	0.00	341.00	0.00	249

D.E., Desviación estandar

Fuente: Red de Estaciones Meteorológicas de la Agroindustria Azucarera Guatemalteca, Administrada por el Instituto Privado de Investigación en Cambio Climático, ICC, (2017). NA=No Aplica.

3.2 MATERIAL VEGETAL

Todos los genotipos de caña de azúcar evaluados, para efectos de este trabajo, son producto del trabajo de mejora genética del Programa de Variedades de CENGICANA, excepto los cultivares testigo. Los genotipos en evaluación son producto de recombinaciones genéticas controladas según los progenitores correspondientes (Anexo 1).

El número de genotipos “CG” estudiados en cada experimento varió entre 112 para el experimento de genotipos no florecedores a 134 para los experimentos tanto florecedores

como no florecedores del estrato litoral (Tabla 3.2). El cultivar CP72-2086 se incluyó como testigo común en todos los experimentos sin importar el estrato ni la época de evaluación porque se trata de un cultivar de amplia difusión en Guatemala, y ocupa cerca de un 36% del total del área de cultivo de la caña de azúcar en Guatemala (Orozco y Buc 2016). Este cultivar es el testigo obligatorio en todos los experimentos de mejora genética de CENGICANA en Guatemala. Adicionalmente se incluyeron otros testigos, que son variedades de importancia relativa en cada una de las áreas y épocas de evaluación. Dentro de estos se incluyen dos testigos de variedades guatemaltecas (CG98-78 y CG98-46), la cuales, actualmente, tienen áreas importantes de cultivo en Guatemala.

El nombre de los genotipos guatemaltecos en evaluación (Anexo 1) tienen el siguiente significado: consideremos el genotipo “**CG12-534048**”; CG, significa CENGICANA-Guatemala; el siguiente número, **12**, corresponde al año de su selección en el Estado I del proceso de mejoramiento. Los siguientes tres dígitos corresponden al número de cruce de la cual el genotipo es una progenie. Finalmente, los últimos tres dígitos corresponden con el número de selección correlativa dentro de la cruce correspondiente.

Tabla 3.2. Número de genotipos en evaluación por experimento.

Experimento	Número de Genotipos CG	Cultivares testigos	Progenitores de los testigos
Estrato medio_flor	133	CP72-2086	(CP62-374 X CP63-588)
		CG98-78	(Tuc68-19 X CP57-603)
		CP88-1165	(CL61-620 X CP81-1302)
Estrato litoral_flor	134	CG98-46	(CP56-59 X CP57-603)
		CP72-2086	(CP62-374 X CP63-588)
		CP73-1547	(CP66-1043 X CP56-63)
Estrato medio_no_flor	134	CP88-1165	(CL61-620 X CP81-1302)
		CG98-78	(Tuc68-19 X CP57-603)
		CP72-2086	(CP62-374 X CP63-588)
Estrato litoral_no_flor	112	CP73-1547	(CP66-1043 X CP56-63)
		SP79-1287	(? x ?)
		CG98-10	(CL61-620 X CP81-1302)
		CP88-1165	(CP62-374 X CP63-588)
		CP72-2086	(? x ?)
		SP71-6161	

? x ? = Progenitores desconocidos

Algunos genotipos en el listado del Anexo 1 poseen nombres mucho más cortos. Esto se debe a que se acortó el nombre, ya que estos genotipos avanzaron (en ciclos posteriores a la realización de este estudio) a una siguiente etapa de evaluación (el Estado III-Incremento). El nombre se redujo con el objetivo de hacerlos más accesibles en su identificación en posteriores evaluaciones y las actividades posteriores de promoción. Todo lo anterior corresponde con el procedimiento general de trabajo del Programa de Variedades de CENGICAÑA, en cuanto al manejo de su sistema y nomenclaturas. Para efectos de trazabilidad, los cambios de nombre efectuados se consignaron en todos los libros de campo donde el genotipo figurara.

3.3 EVALUACIÓN Y SELECCIÓN

Las evaluaciones se realizaron en los cuatro experimentos durante dos años de cultivo. El primer año se suele llamar en Guatemala “plantía”, en otros países se le denomina “caña-planta”, “Plantilla” o simplemente “planta” (Larrahondo, 1995). El segundo año de evaluación, después de la cosecha (corte) del primer año, corresponde al primer rebrote de la caña cortada, el cual se denomina “**primera soca**” (Orozco *et al.*, 2012). En este sentido se hicieron las evaluaciones en los cuatro experimentos, obteniendo siete conjuntos de datos, sin poder tomar en cuenta uno de los conjuntos perdidos por la quema mal intencionada antes mencionada.

De los cuatro experimentos, dos fueron denominados del tipo florecedor (**flor**) y los otros dos del tipo no florecedor (**no_flor**). Los genotipos **flor** fueron evaluados en los meses de diciembre de 2014 y 2015, respectivamente, para los ciclos de plantía y 1ª soca; en tanto que los genotipos **no_flor**, fueron evaluados en los meses de marzo de 2015 y 2016, respectivamente, para el ciclo del plantía y el ciclo de 1ª soca.

Se hicieron cuatro rondas de evaluación para cada uno de los siete conjuntos de datos:

- a. Un mes después de plantados los experimentos se hizo una evaluación sobre las condiciones de emergencia, malezas y riego apropiado.
- b. A los cinco meses después de la plantación o del rebrote se realizó una evaluación general agronómica por escalas en cuanto a la población, altura, y diámetro. En ese

mismo momento se realizó una evaluación por plagas y enfermedades de todos los genotipos en todos los experimentos.

- c. Posteriormente, a los once meses después de la plantación o corte, se realizó un procedimiento de evaluación agronómica. Una vez más se realizó una evaluación por escalas, de la población, altura diámetro, y aspectos fitopatológicos. En este momento, se realizó una evaluación por sólidos solubles (°Brix) en todos los genotipos que, a juicio del fitomejorador, cumplieron con los requisitos agronómicos establecidos para el Programa de Variedades (Tabla 3.3), así como el buen desempeño ante enfermedades (bióticas y abióticas). Los genotipos que presentaran valores superiores del 20% de concentración de sólidos solubles fueron sujetos de una preselección.
- d. Los genotipos preseleccionados según el epígrafe anterior fueron nuevamente evaluados. Esta evaluación se realizó con mucho mayor detalle, tomando todas las medidas que arrojaron los datos necesarios para el cálculo de las variables de respuesta, la cuales se presentan en el acápite de variables en estudio. Se tomó una muestra de quince esquejes de 30 cm de largo para llevarlos a su análisis al laboratorio agronómico de CENGICAÑA, en donde se realizaron todos los análisis de jugos, los cuales también se resumen en el acápite de variables de respuesta.
- e. Una ronda final de evaluación a los 12 meses se realizó con el fin de corroborar las calidades de las evaluaciones y establecer si algún otro genotipo aún podría ser rescatado para un análisis más profundo.

Tabla 3.3. Listado de requisitos del cliente para las variedades que libera CENGICAÑA En Guatemala.

PRODUCTO	REQUISITOS	VARIABLES					
Variedades de Caña de Azúcar	1. Que la variedad CG o introducida en promedio supere en rendimiento de azúcar por unidad de área (TAH) a la variedad CP72-2086, o iguale o supere a otras variedades de interés del cliente.	Mayor t de azúcar por ha TAH que la variedad testigo CP72-2086 y a otras variedades de interés del cliente, que cumpla al menos un componente del TAH (TCH o Pol % caña).					
	2. Que la variedad CG o introducida posea resistencia adecuada a enfermedades.	Carbón	Escaldadura Foliar	Mosaico	Roya Marrón	Roya Naranja	Amarillamiento
		(0-10)	(0-10)	(0-10)	Incidencia/ Severidad (1-50) / (0-9)	Incidencia/ Severidad (1-50) / (0-9)	(0-10)
		≤ 1	≤ 1	≤ 2B	≤ 15 / 5	≤ 10 / 5	≤ 5B
	3. Que la variedad CG o introducida presente buenas características de manejo.	<ul style="list-style-type: none"> - Emergencia y rebrote mejor o igual que CP72 -2086 y a otras variedades de interés del cliente. - Cierre natural mejor o igual que CP72-2086 y a otras variedades de interés del cliente. - Aspecto de la planta mejor o igual que CP72 -2086 y a otras variedades de interés del cliente. - Corcho menor o igual que CP72-2086 y a otras variedades de interés del cliente. - Con o sin aplicaciones de madurante, mayor TAH que las variedades testigo CP72-2086 y a otras variedades de interés del cliente. - t de fibra por ha TFH mayor o igual que CP72-2086 y a otras variedades de interés del cliente. <p>En casos especiales se seleccionarán variedades que cumplan con los requisitos 1 y 2, aunque no cumplan con alguna variable de manejo.</p>					
	4. Que la variedad CG o introducida tenga adaptabilidad específica o general en la zona cañera guatemalteca.	<p>Adaptabilidad específica: Promedio de TAH en al menos una localidad dentro de una zona en al menos dos cortes superior al TAH de la variedad CP72-2086 y a otras variedades de interés del cliente.</p> <p>Adaptabilidad general: promedio de TAH de localidades dentro de una zona en al menos dos cortes superior al TAH de la variedad CP72-2086 y a otras variedades de interés del cliente.</p>					

3.4 DISEÑO EXPERIMENTAL

Todos los ensayos fueron establecidos en un esquema para un Diseño de Bloques Incompletos (Montgomery, 1991). Este diseño es conocido en algunos ámbitos de investigación como Alfa-Latice. Debido a la dinámica de evaluación y selección, finalmente todos los conjuntos de datos fueron analizados para un Diseño de Bloques Incompletos, Parcialmente Balanceado (Montgomery, 1991).

3.5 VARIABLES DE ESTUDIO

Se realizaron un total de 58 mediciones diferentes (Anexo 2). Con estas mediciones se definieron catorce variables de respuesta. Se computaron también variables de clase, con el objeto de deslindar los componentes ambientales y genéticos. Las variables empleadas se describen en la Tabla 3.4. Todas las variables numéricas fueron empleadas como variables de respuesta.

Tabla 3.4. Variables evaluadas, correspondencia y tipo de variable en R.

No.	Nombre de la Variable	Código	Dimensional	Nombre de la Variable en R	Tipo de variable en R
1	Conjunto de Datos	NA	NA	TRIAL=	as.factor(TRIAL)
2	Repetición	NA	NA	REP=	as.factor(REP)
3	Bloque incompleto	NA	NA	BLOCK=	as.factor(Block)
4	Nombre del Genotipo	NA	NA	ENTRY=	as.factor(ENTRY)
5	°Brix medido en el campo	brixprom_1	%	BRIX_CAMPO=	as.numeric(brixprom_1)
6	Peso de una muestra de cinco tallos de caña	peso5stalk_2	kg	PESO=	as.numeric(peso5stalk_2)
7	Oquedad leve	oqueleve_3	%	OQUEDAD1=	as.numeric(oqueleve_3)
8	Oquedad intermedia	oqueinter_4	%	OQUEDAD2=	as.numeric(oqueinter_4)
9	Oquedad severa	oqueseve_5	%	OQUEDAD3=	as.numeric(oqueseve_5)
10	Pol%Caña	pol_6	%	POL=	as.numeric(pol_6)
11	°Brix del jugo medido en el laboratorio	brixjugo_7	%	BRIX_JUGO=	as.numeric(brixjugo_7)
12	Pureza del Jugo, determinada en el laboratorio	purity_8	%	PUREZA=	as.numeric(purity_8)
13	Población de tallos	pop_9	Frecuencia absoluta por metro lineal	POBLACION=	as.numeric(pop_9)
14	Altura de corte de la caña	alt_10	m	ALTURA=	as.numeric(alt_10)
15	Diámetro medio de la caña	diam_11	mm	DIAMETRO=	as.numeric(diam_11)
16	Proporción de tallos florecidos	flower_12	%	FLORACION=	as.numeric(flower_12)
17	Corcho o tejido interno seco en las cañas	pith_13	%	CORCHO=	as.numeric(pith_13)
18	Materia Seca de la muestra	fiber_14	%	FIBRA=	as.numeric(fiber_14)
19	TAH	NA	t de azúcar por ha	TAH=	as.numeric(TAH_15)
20	Componente Principal 1	NA	Adimensional	CP1=	as.numeric(CP1)
21	Componente Principal 2	NA	Adimensional	CP2=	as.numeric(CP2)
22	Componente Principal 3	NA	Adimensional	CP3=	as.numeric(CP3)

NA: No Aplica.

Para las variables, Oquedad leve (7), Oquedad intermedia (8), Oquedad severa (9), proporción de tallos florecidos (16), Corcho o tejido interno seco en las cañas (17), un valor bajo de estas variables, es considerado como un valor bueno. Es decir, lo contrario que para el resto de las variables, donde un valor alto es considerado como un valor bueno. Para que todas las variables se consideraran en el sentido de que, un valor alto es un valor bueno, las variables mencionadas, fueron transformadas para que un valor alto de éstas también fuera considerado, efectivamente, como un valor bueno.

Para poder reorientar en el sentido positivo la dirección de éstas, se empleó la siguiente transformación:

$$\text{valor positivo de la variable} = \text{valor máximo de la variable} - \text{el valor de la observación.}$$

Así, un valor alto es aquel que se aleja de un valor considerado como malo, por tanto, está en la misma dirección del resto de variables, en las cuales, un valor alto es considerado un valor bueno.

3.6 ANÁLISIS DE LA INFORMACIÓN

En donde sea de aplicación, los códigos para las significaciones estadísticas se seguirán según el siguiente planteamiento:

****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

3.6.1 Análisis general por variables

Se realizaron una serie de análisis individuales de las principales variables de respuesta, 14 en total. Las principales medidas empleadas fueron de orden estadístico:

- Análisis de frecuencias para construir Histogramas (Spiegel, 1987).
- Medias aritméticas y sus correspondientes desviaciones (Spiegel, 1987).
- Pruebas de Ji-cuadrado (χ^2) (Steel y Torrie, 1986).
- Pruebas de t de Student (Montgomery, 1991)
- Análisis de correlación entre todas las variables mediante el cálculo de los coeficientes por el método de mínimos cuadrados de Pearson (Mendenhal y Beaver, 1991). Los cálculos del coeficiente de correlación y el nivel de significación de los modelos se realizó mediante el uso del paquete Hmisc (Harrell *et al.*, 2017). Para cada variable en particular se generaron trece regresiones.

El modelo lineal ajustado fue:

$$y_i = \beta_0 + \beta_1 x_i \quad 3.1$$

Donde,

y_i = la variable dependiente (cualquiera de las catorce), donde $i = 1, 2, \dots, 14$

β_0 = el coeficiente intercepto

β_1 = el coeficiente pendiente

x_i = la variable independiente, donde $i = 1, 2, \dots, 14$

Las figuras y gráficos relacionados se realizaron con el programa R (R Core Team, 2017) mediante el uso de los paquetes **corrplot** (Wei y Simko, 2016), **ggbiplot** (Wickham, 2009) y **GGally** (Scholerke *et al.*, 2016).

3.6.2 ANÁLISIS GENERAL DE VARIACIÓN COMBINADO

Se realizó un Análisis General Combinado para la variable TAH siguiendo el modelo lineal matemático que se presenta en la Ecuación 3.2. En un primer análisis, todos los efectos fueron considerados fijos, por lo que la estimación de cada factor se realizó por medio de (BLUE) (Piepho *et al.*, 2008). En un segundo análisis, se obtuvieron resultados para un modelo lineal mixto (Bates *et al.*, 2015); en este segundo análisis, todos los efectos, excepto el relacionado con los genotipos de caña (ENTRY), fueron considerados aleatorios. En este último análisis se incluyeron en el análisis los *scores* de los primeros cuatro CPs; así mismo, para los efectos aleatorios se utilizó el método estimación basado en REML/BLUP (Piepho *et al.*, 2008; Thompson, 2008; Robinson, 1991).

$$y_{ijklmn} = \mu + \gamma_i + \delta_j + \eta_k + \rho_l + \eta_m + \gamma\delta_{ij} + \gamma\eta_{ik} + \rho\theta_{ln} + \varepsilon_{ijklmn} \quad 3.2$$

Donde,

y = la variable dependiente (TAH)

μ = la media general

γ = el efecto del i -ésimo ciclo de evaluación. $i = 1, 2$

δ = el efecto del j -ésimo estrato, $j = 1, 2$

η = el efecto del k -ésimo nivel del patrón de floración, $k = 1, 2$

ρ = el efecto de la l -ésima repetición, $l = 1, 2$

η = el efecto del m -ésimo genotipo, $m = 1, 2, \dots, m$ (varía según el conjunto de datos)

θ = el efecto del n -ésimo bloque incompleto, $n = 1, \dots, n$ (varía según el conjunto de datos)

ε = los residuos asociados al resto de factores.

El modelo incluye una serie de interacciones: Entre Ciclos de cultivo (γ) con Estratos (δ), Ciclos (γ) con el patrón de floración (η) y la interacción entre repeticiones (ρ) y el bloque incompleto (θ). Las interacciones con respecto a los factores ambientales y los genotipos, no fueron incluidas en el modelo por cuanto las evaluaciones según ambiente, contuvieron materiales diferentes en cada conjunto de datos.

3.6.3 Heredabilidad

El elemento de ponderación sugerido para varios índices de selección es el peso económico relativo que pueda tener una variable de valoración en particular (Smith, 1936; Kempthorne y Nordskog, 1959; Cerón-Rojas *et al.*, 2006). Sin embargo, la estimación de estos ponderadores es complicada y requiere una gran cantidad de información, que en sí se convierten en campos específicos de estudio. Como un sustituto válido de los ponderadores económicos se pueden emplear los valores de heredabilidad (Williams, 1962; Pérez-Elizalde *et al.*, 2014).

La caña de azúcar es un cultivo de reproducción vegetativa y esta es la principal forma de propagación a través de todo el proceso de selección en el programa de mejoramiento genético de CENGICAÑA (Orozco *et al.*, 2012). En general, las poblaciones bajo estudio son progenies segregantes, provenientes del estado de selección inmediato anterior. Por otro lado, se tiene el conocimiento de los progenitores de cada genotipo, sin embargo, el número de éstos es muy alto, y la caracterización correspondiente no existe. Hacer las caracterizaciones en las condiciones de selección actuales resultaría en experimentos demasiado grandes y la labor sería inmensa. Estas circunstancias condujeron al cálculo de la *heredabilidad en el sentido amplio* (*Broad sense heritability*) (H^2) (Fehr, 1987; Skinner *et al.*, 1987; Holland *et al.*, 2003), en detrimento del cálculo de la *heredabilidad en el sentido estricto* (*narrow sense heritability*) (h^2) (Fehr, 1987; Skinner *et al.*, 1987; Holland *et al.*, 2003).

El cálculo de los valores de heredabilidad en el sentido amplio (H^2), por carácter, se hizo sobre la base de las medias de las variedades (*entry-mean basis*) según el modelo que plantea la Ecuación 3.3. (Fehr, 1987):

$$H^2 = \frac{\sigma_g^2}{\frac{\sigma_e^2}{rt} + \frac{\sigma_{ge}^2}{t} + \sigma_g^2} \quad 3.3$$

Donde,

H^2 = la heredabilidad en el sentido amplio

σ_g^2 = la varianza genética

σ_e^2 = el error Experimental

σ_{ge}^2 : = la varianza de la interacción del genotipo por el ambiente

r = el número de repeticiones

t = el número de ambientes de evaluación

La fuente de variación para los ambientes lo constituyeron los dos ciclos (años) de evaluación: plantía y 1ª soca. La Ecuación 3.3 se aplicó en todos los experimentos, excepto en el experimento 3, el cual contaba con información de solamente un año por lo que en éste se empleó la siguiente expresión (Fehr, 1987):

$$H^2 = \frac{\sigma_g^2}{\frac{\sigma_e^2}{rt} + \sigma_g^2} \quad 3.4$$

El significado de los términos es el mismo que se presenta en la Ecuación 3.3.

Los valores de las varianzas respectivas fueron obtenidos mediante el análisis de variación de la regresión de los modelos respectivos. Lo anterior se realizó en un esquema de Modelos Lineales Mixtos, en donde todas las variables del modelo fueron consideradas como variables aleatorias. El análisis se efectuó mediante el empleo del paquete para análisis estadísticos en R **lme4** (Bates *et al.*, 2015). Con los valores obtenidos de las varianzas correspondientes se procedió al cálculo de los valores de heredabilidad, según las ecuaciones 3.3 y 3.4.

Para poder incorporar el efecto del ambiente en el cálculo de los valores de heredabilidad, se utilizaron los datos combinados de cuatro grupos de datos, agrupados según cada uno de los experimentos, los cuales fueron evaluados durante dos años. Para tal efecto, se ajustó el modelo lineal de la Ecuación 3.5, mediante el uso del comando **lmer** del paquete **lme4** (Bates *et al.*, 2015).

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + \beta\delta_{jl} + \tau\gamma_{ik} + \varepsilon_{ijkl} \quad 3.5$$

Donde,

y_{ijkl} = La variable de respuesta asociada a la i-ésima variedad, en la j-ésima repetición, en el k-ésimo año, en el l-ésimo bloque incompleto

τ_i = el i-ésimo genotipo (Componente genético)

β_j = la j-ésima repetición (Componente ambiental)

γ_k = el k-ésimo ciclo de evaluación (Componente ambiental)

$\beta\delta_{jl}$ = el efecto del l-ésimo bloque incompleto en la j-ésima repetición (Componente ambiental)

$\tau\gamma_{ik}$ = el efecto de la interacción entre el i-ésimo genotipo con el k-ésimo ciclo de evaluación (Componente ambiental)

ε_{ijkl} = los residuos (Error experimental)

Como el modelo ajustado en realidad fue para un modelo lineal mixto, en donde todos los términos de 3.4 fueron considerados aleatorios, la forma matricial del mismo se presenta en la Ecuación 3.6.

$$\mathbf{y} = \mathbf{Z}_{u_1} + \mathbf{Z}_{u_2} + \mathbf{Z}_{u_3} + \mathbf{Z}_{u_4} + \mathbf{Z}_{u_5} + \boldsymbol{\epsilon} \quad 3.6$$

Donde,

\mathbf{y} = el vector de respuestas y_n para cada variable en estudio

\mathbf{Z} = la matriz incidente para los efectos aleatorios

u_1 = el vector del efectos aleatorios de τ_i

u_2 = el vector de efectos aleatorios de β_j

u_3 = el vector de efectos aleatorios de γ_k

- u_4 = el vector de efectos aleatorios de $\beta\delta_{jl}$
 u_5 = el vector de efectos aleatorios de $\tau\gamma_{ik}$
 ϵ = el vector de efectos aleatorios de los residuos

La varianza de la respuesta, $Var(y)$, es reconocida como parte de la sección aleatoria del modelo, de tal forma que:

$$Var(y) = Var(Zu + \epsilon) = ZGZ + R = V \quad 3.7$$

Con,

$$u \sim MVN(u, G)$$

$$\epsilon \sim MVN(u, R)$$

(Ver sección Modelos Lineales Mixtos, BLUE y BLUP, epígrafe. 1.5.6.2.).

Debido a que no hay efectos fijos en el modelo, este no es, estrictamente hablando, un modelo mixto. En todo caso se analiza como tal. El principal producto de este ajuste fueron las varianzas de cada uno de los términos del modelo en 3.5 para las 14 variables listadas en la Tabla 3.4. Estos valores se emplearon para el cálculo de la heredabilidad según las ecuaciones 3.2 y 3.3.

El modelo simplificado en el paquete **lme4** en R fue:

$$\text{Respuesta} \sim (1 | \text{ENTRY}) + (1 | \text{REP}) + (1 | \text{CICLO}) + (1 | \text{BLOCK:REP}) + (1 | \text{ENTRY:CICLO})$$

El ajuste del modelo se realizó para 14 variables que se presentan en la Tabla 3.5.

Tabla 3.5. Listado de variables para las cuales se estimó la heredabilidad y el cálculo de los índices de selección.

No.	Variable.	Código
1	°Brix medido en el Campo (%)	brixprom_1
2	Peso De Cinco tallos (%)	peso5stalk_2
3	Oquedad Leve (%)	oqueleve_3
4	Oquedad Intermedia (%)	oqueinter_4
5	Oquedad Severa (%)	oqueseve_5
6	Pol (%)	pol_6
7	°Brix Jugo (%)	brixjugo_7
8	Pureza del jugo de una muestra (%)	purity_8
9	Frecuencia de la Población por metro lineal	pop_9
10	ALTURA (m)	alt_10
11	DIÁMETRO (mm)	diam_11
12	Flor (%)	flower_12
13	Corcho (%)	pith_13
14	Fibra (%)	fiber_14

3.6.4 Análisis De Componentes Principales (ACP)

Se realizaron varios cómputos de CPs con el objeto de establecer la estructura de los datos y las variables (Hernández, 1998; Jolliffe, 2002). Se computaron CPs para los siete conjuntos de datos (el octavo fue perdido por quema provocada). Se presentan en esta memoria solamente los resultados de los conjuntos de datos **medio_flor_ plantía**, **medio_no_flor_1ª_soca** y **litoral_no_flor_ plantía**.

Las “cargas” (*eigenvectors*, o *loadings*) son los coeficientes de las funciones lineales (CP) y a partir de ellas y de las lecturas de cada observación para cada variable, se generó un *score*. Las CPs se utilizaron como variables de respuesta en otro de análisis de tal manera que el cómputo de las mismas sirvió para la obtención de *scores* según los modelos lineales de las CPs. Estos *scores* fueron empleados como índices de desempeño (selección). Las “cargas” indican que si éstos están asociados a determinadas variables originales dirigirán la orientación de la variable, así, si el valor de estas cargas es negativa y aportan de forma importante a la CP, entonces, cuanto más bajo sea el *score*, mejor es el desempeño con respecto a las variables que estén influyendo en la CP. Si el signo es positivo, entonces cuanto más alto sea el *score*, mejor es el desempeño de la observación individual.

Todos los ACPs fueron realizados en el programa para análisis estadístico R (R Core Team, 2017) mediante la aplicación de la función **prcomp** del paquete **stats** (R Core Team, 2017). Para el cómputo de los ACP's se utilizaron las 14 variables listadas en la Tabla 3.5.

Las figuras relacionadas con los ACP's fueron construidas a partir de instrucciones realizadas con los paquetes **lattice** (Deepayan, 2008), **rgl** (Adler *et al.*, 2017) y **devtools** (Wickham y Chang, 2017).

3.6.5 Análisis De Modelos Lineales Mixtos Para Los Scores De Las CPs

Jolliffe, (2002) indica que las CPs son definidas como una transformación “ortonormal” (*orthonormal*) de x , el cual es el vector de las p variables aleatorias. En este sentido, cada CP es una variable que puede ser sujeta de otros tipos de análisis como el de conglomerados (*cluster analysis*), y otros análisis de orden paramétrico.

Considerando a los *scores* de las tres primeras CPs como variables de respuesta, y haciendo un ejercicio de éstos como índices de desempeño, se realizaron varios Análisis de Varianza (ANOVA) con el objeto de buscar la diferenciación entre genotipos a partir de los CPs. El modelo lineal ajustado para cada conjunto de datos se presenta en la Ecuación 3.8.

$$y_{ijk} = \mu + \tau_i + \gamma_j + \gamma\delta_{jk} + \varepsilon_{ijk} \quad 3.8$$

Donde,

y_{ijk}	=	la variable de respuesta (CP1, CP2 y CP3)
μ	=	el efecto de la media
τ_i	=	el efecto del i-ésimo genotipo en evaluación (El efecto genético)
γ_j	=	el efecto de la j-ésima repetición
$\gamma\delta_{jk}$	=	el efecto de la interacción de la j-ésima repetición con el k-ésimo bloque incompleto
ε_{ijk}	=	el efecto de los residuos (error experimental)

En este modelo solamente el efecto genético (τ_i) se consideró fijo, el resto de los términos se consideraron efectos aleatorios. Lo anterior se realizó porque el número de bloques incompletos en cada experimento no estaba equilibrado, por lo que éstos, conjuntamente con las repeticiones, se analizaron como efectos aleatorios. En su forma matricial para el análisis de modelos mixtos el modelo matemático queda según se presenta en la Ecuación 3.9.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_0\mathbf{c} + \mathbf{Z}_1\mathbf{g} + \boldsymbol{\epsilon} \quad 3.9$$

Donde,

\mathbf{y} = el vector $p \times 1$ de respuestas y (el clon en una localidad) para cada variable en estudio (PC1, PC2, PC3)

\mathbf{X} = la matriz incidente de 1's y 0's para los efectos fijos (genotipos)

\mathbf{Z} = la matriz incidente para los efectos aleatorios

$\boldsymbol{\beta}$ = el vector $i \times 1$ de los efectos fijos por el desempeño de los genotipos

\mathbf{c} = el vector $j \times 1$ del efecto aleatorio de las repeticiones γ_j

\mathbf{g} = el vector $j \times k$ del efectos aleatorio producto de la interacción $\gamma\delta_{jk}$

$\boldsymbol{\epsilon}$ = el vector $p \times 1$ de efectos aleatorios de los residuos

El ajuste del modelo lineal mixto se realizó mediante la aplicación de la función **lmer** del paquete **lme4** (Bates *et al.*, 2015). Los gráficos de cuantil-cuantil (QQ) se realizaron mediante el uso del paquete **lattice** (Deepayan, 2008). La comparación de promedios se realizó mediante el ajuste de Tukey mediante la aplicación del paquete para R **predictmeans** (Luo *et al.*, 2014).

3.6.6 Análisis Multidimensional De Variación (MANOVA)

Para contar con una herramienta que permitiera saber si las variaciones en el desempeño de los genotipos era función de todas las catorce variables en estudio juntas, se realizó un Análisis Multidimensional de Variación (MANOVA) (Rencher, 2002). Esta información fue útil para establecer cuál fue el nivel de relación funcional entre las 14 variables dependientes y las variables independientes en el modelo (Tabla 3.1). Por lo anterior, se ajustaron los modelos para los siete conjuntos de datos para un Análisis Multidimensional de Variación (MANOVA). El modelo para el ajuste se presenta en la Ecuación 3.10. El modelo es muy similar al presentado en 3.9. Por esta razón, y para diferenciarlos, se cambió la nomenclatura de las variables. La Ecuación 3.10, en notación matricial, presenta el modelo ajustado. Hay que señalar que el modelo, si bien es en notación matricial, **no** es un modelo mixto.

$$\mathbf{V}_t = \mathbf{V}_p + \mathbf{V}_m + \mathbf{V}_{mq} + \mathbf{V}_e \quad 3.10$$

Donde,

V_t = el vector $t \times 1$ de las respuestas para las t variables (14)

V_p = el efecto de los genotipos en una matriz $p \times p$.

V_m = el efecto de las repeticiones en una matriz $m \times m$.

V_{mq} = el efecto de la interacción entre las repeticiones y los bloques incompletos en una matriz $m \times q$.

Los análisis fueron realizados mediante la aplicación de los paquetes para **R**, **MCMCglmm** (Hadfield, 2010), **lme4** (Bates *et al.*, 2015), **reshape2** (Wickham, 2007), e **lmerTest** (Kuznetsova *et al.*, 2016).

3.6.7 Los Índices de Selección

3.6.7.1 Función Discriminante Propuesta Por Smith (Smith, 1936)

El índice “Smith” se computó según la propuesta teórica de Smith (1936). Los detalles teóricos se presentan en la sección de marco bibliográfico referencial, en Smith (1936) y en Cerón-Rojas *et al.*, (2006). La computación de los valores genéticos del índice, cuyos resultados completos se presentan en el Anexo 3, se realizó mediante el uso del paquete para **R**, **Rindsel** (Pérez-Elizalde *et al.*, 2014). El paquete **Rindsel**, requiere del uso de otros paquetes de R: **lme4** (Bates *et al.*, 2015), **MASS** (Venables y Ripley, 2002) y **Hmisc** (Harrell *et al.*, 2017). Este índice requiere que se apliquen ponderadores económicos. Al no contar con ellos, se utilizaron los valores de heredabilidad, cuyo procedimiento de estimación se presenta en el acápite 4.6.3. Los valores de heredabilidad empleados en esta sección se estimaron según la procedencia de cada conjunto de datos. Así, se contó con cuatro grupos de valores de heredabilidad: **medio_flor**, **medio_no_flor**, **litoral_flor**, **litoral_no_flor**. Según fuera la procedencia del conjunto de datos, se le aplicó el valor de heredabilidad correspondiente, por ejemplo, al conjunto de datos **medio_flor_plantía** se le aplicaron los valores de heredabilidad del grupo **medio_flor**.

Para la presentación de resultados, se empleó el criterio de selección del 10% de la fracción superior de los genotipos según el valor de los índices Smith y KN. El programa **Rindsel** ofrece tres opciones 5, 10 y 20%, aunque este valor es siempre decisión del fitomejorador. Lo más común es el empleo de la selección del 10% de la fracción superior

según la variable de evaluación, en este caso, índices (Pérez-Elizalde *et al.*, 2014; Pesek y Baker, 1970).

3.6.7.2 Índice Restrictivo De Kempthorne Y Nordskog, KN (Kempthorne Y Nordskog, 1959)

Al igual que en la función discriminante de Smith (Smith, 1936), este índice también requiere de ponderadores económicos. De similar forma, en el cómputo de este índice se utilizaron los valores de heredabilidad como ponderadores. El índice fue computado mediante la aplicación del paquete para **R**, **Rindsel** (Pérez-Elizalde *et al.*, 2014). El paquete **Rindsel** requiere del uso de otros paquetes de R: **lme4** (Bates *et al.*, 2015), **MASS** (Venables y Ripley, 2002), y **Hmisc** (Harrell *et al.*, 2017). En virtud de que es un índice restrictivo, se realizó un procedimiento de codificación para establecer la restricción. El código “1” correspondió a las variables que se mantenían sin cambios (no ganancia genética), en tanto que “0” se utilizó para codificar las variables de las cuales se requieren cambios, en términos de ganancia genética (cambios), ya sea esta positiva o negativa (Kempthorne y Nordskog, 1959; Cerón-Rojas *et al.*, 2008a).

Las variables a las cuales se les asignó un valor de “cambio” (“0”) fueron: **brixprom_1**, **peso5stalk_2**, **pol_6**, **pith_13**, **fiber_14**. A las variables a las que se les asignó código para “no cambio” (“1”) fueron: **oqueleve_3**; **oqueinter_4**; **oqueseve_5**; **brixjugo_7**; **purity_8**; **pop_9**; **alt_10**, **diam_11**; **flower_12** (Tabla 3.5).

La decisión de la asignación de la restricción se fundamentó en la experiencia y en la competencia de los fitomejoradores del Programa de Variedades Programa de Variedades de CENGICANÑA.

Para la presentación de resultados, se empleó el criterio de selección del 10% de la fracción superior de los genotipos según el valor de los índices Smith y KN. El programa **Rindsel** ofrece tres opciones 5, 10 y 20%, aunque este valor es siempre decisión del fitomejorador. Lo más común es el empleo de la selección del 10% de la fracción superior según la variable de evaluación, en este caso, índices (Pérez-Elizalde *et al.*, 2014; Pesek y Baker, 1970).

3.6.7.3 *Eigenanalysis-Based Selection Index Method (ESIM)*

Este índice se calcula utilizando como base un ACP, discutido en el epígrafe 3.6.3. El índice fue computado mediante la aplicación del paquete para **R**, **Rindsel** (Pérez-Elizalde *et al.*, 2014). El paquete **Rindsel** requiere del uso de otros paquetes de **R**: **lme4** (Bates *et al.*, 2015), **MASS** (Venables y Ripley, 2002), y **Hmisc** (Harrell *et al.*, 2017). La presentación y discusión de resultados, al igual que en Smith y KN, se hicieron sobre la base de la selección del 10% de la fracción superior de los genotipos en evaluación.

3.6.7.4 *Restrictive Eigenanalysis-Based Selection Index Method (RESIM)*

Este índice se calcula utilizando como base un ACP. El índice fue computado mediante la aplicación del paquete para **R**, **Rindsel** (Pérez-Elizalde *et al.*, 2014). El paquete **Rindsel** requiere del uso de otros paquetes de **R**: **lme4** (Bates *et al.*, 2015), **MASS** (Venables y Ripley, 2002), y **Hmisc** (Harrell *et al.*, 2017). Adicionalmente, es preciso restringir el análisis a variables de interés. Al igual que en el índice KN, se realizó un procedimiento de codificación para establecer la restricción. El código “1” es para las variables que se mantendrán sin cambios (no ganancia genética), en tanto que “0” es para codificar las variables de las cuales se requieren cambios (Kempthorne y Nordskog, 1959; Cerón-Rojas *et al.*, 2008a). La codificación que se empleó en el cómputo de este índice fue el mismo que para el índice KN.

La decisión de la asignación de la restricción se fundamentó en la experiencia y en la competencia de los fitomejoradores del Programa de Variedades de CENGICAÑA.

La presentación y discusión de resultados, se hicieron sobre la base de la selección del 10% de la fracción superior de los genotipos en evaluación.

3.6.7.5 *CPs Como Índices De Desempeño*

En un análisis de ACP (epígrafe 1.2.3), se obtuvo información sobre los *eigenvalues*, *eigenvectors* y los *scores* para cada observación. Los valores de los *scores* se computaron como índices de desempeño y se anexaron a la información relativa a cada observación (Anexo 3). Los procedimientos de cómputo ya se presentaron en el epígrafe 3.6.3.

La selección final de los genotipos, mediante el uso de un ACP se realizó de manera anidada de la siguiente manera:

- a. Se realizó un análisis de varianza a los *scores* de la CP1.
- b. Según los signos de valores sobresalientes de los *eigenvectors* se ordenaron los genotipos en orden descendente según el desempeño.
- c. Se eligieron los mejores genotipos de acuerdo con esta CP, de tal forma que alcanzaran el 25% de la fracción superior. El 25% es arbitrario y se hizo con el fin de contar con más genotipos que pudieran ser evaluados posteriormente, según las CP1 y CP2.
- d. Seguidamente, se evaluaron los genotipos según la CP2 y se marcaron los de mejor desempeño.
- e. A los genotipos marcados como los mejores en las CP1 y CP2, se les evaluó según los *scores* de la CP3.
- f. Se seleccionó finalmente la fracción superior del 10% después de las evaluaciones consecutivas anidadas.
- g. Para la presentación de resultados, se empleó el criterio de selección del 10% de la fracción superior de los genotipos según el valor de los índices Smith y KN. El programa **Rindsel** ofrece tres opciones 5, 10 y 20%, aunque este valor es siempre decisión del fitomejorador. Lo más común es el empleo de la selección del 10% de la fracción superior según la variable de evaluación, en este caso, índices (Pérez-Elizalde *et al.*, 2014; Pesek y Baker, 1970).

3.6.7.6 TAH, Índice t De Azúcar Por ha

La estimación del **TAH** se computó como una operación aritmética, la que se presenta en la Ecuación 3.11.

$$\text{TAH} = \left[\left(\left(A \times \left[\frac{B}{5} \right] \times 10000 \right) / 30 \right) / 1000 * C \right] / 100 \quad 3.11$$

Donde,

A = Número de tallos por unidad experimental

B = Peso de una muestra de cinco tallos en kg

C = Concentración de azúcar en % (Pol % caña)

El valor 10000 se utiliza para convertir el dato a ha, ya que está en m²

El valor 30 se refiere al área de una unidad experimental, ya que está en m²

El valor 1000 se utiliza para convertir la respuesta final a t por ha

Dado que Pol % caña está dado en porcentaje, es preciso convertirlo a unidades relativas sobre “1”, de ahí la necesidad de dividir el resultado por cien.

El TAH fue empleado, indistintamente, como una variable de respuesta común o como un índice de desempeño. Al observar la forma del cálculo se aprecia que las variables más importantes en término de rendimientos están involucradas en el cálculo del TAH, por lo que se puede considerar como un valor que agrupa los efectos de otras variables, es decir, un índice.

Para la presentación de resultados, se empleó el criterio de selección del 10% de la fracción superior de los genotipos según el valor de los índices Smith y KN. El programa **Rindsel** ofrece tres opciones 5, 10 y 20%, aunque este valor es siempre decisión del fitomejorador. Lo más común es el empleo de la selección del 10% de la fracción superior según la variable de evaluación, en este caso, índices (Pérez-Elizalde *et al.*, 2014; Pesek y Baker, 1970).

3.6.7.7 BLUP (*Best Linear Unbiased Prediction*). El Mejor Estimador Lineal Insesgado

BLUP se estimó mediante la aplicación de los procedimientos de la solución de las ecuaciones de un modelo lineal mixto (Henderson *et al.*, 1959; Robinson, 1991; Piepho *et al.*, 2008; Thompson, 2008). BLUP se calculó utilizando como base la variable TAH. Para el cálculo de este predictor se utilizó también el paquete para análisis estadísticos en R **lme4** (Bates *et al.*, 2015). En este caso la variable de respuesta analizada fue TAH y el modelo lineal utilizado se aplicó para cada uno de los siete conjuntos de datos, por lo que las variables localidad y años ya no fueron incluidas en el modelo. Lo anterior se realizó de esta manera porque las selecciones de materiales genéticos fueron diferenciales. El modelo en cuestión se presenta en su forma lineal en la Ecuación 3.12.

$$y_{ijk} = \mu + \tau_i + \gamma_j + \gamma\delta_{jk} + \varepsilon_{ijk} \quad 3.12$$

Donde,

- y_{ijk} = la variable de respuesta (TAH)
 μ = el efecto de la media
 τ_i = el efecto del i -ésimo genotipo en evaluación (el efecto genético)
 γ_j = el efecto de la j -ésima repetición
 $\gamma\delta_{jk}$ = el efecto de la interacción de la j -ésima repetición con el k -ésimo bloque incompleto
 ε_{ijk} = el efecto de los residuos (error experimental)

Hay que señalar que el modelo en la expresión 3.12 es el mismo que se presentó en la Ecuación 3.8, la diferencia la hacen las variables de respuesta, en este caso es el TAH. Debido a que en este caso todos los términos de 3.12 se consideraron como efectos aleatorios, el mismo modelo lineal para un modelo lineal mixto se presenta en su notación matricial en 3.13.

En su forma matricial, y utilizando como referencia la Ecuación 3.9, es como sigue:

$$\mathbf{y} = \mathbf{Z}_0\mathbf{b} + \mathbf{Z}_1\mathbf{c} + \mathbf{Z}_2\mathbf{g} + \boldsymbol{\epsilon} \quad 3.13$$

Donde,

- \mathbf{y} = el vector $p \times 1$ de respuestas \mathbf{y} para la variable en estudio (TAH)
 \mathbf{Z}_0 = la matriz incidente para los efectos aleatorios de los genotipos
 \mathbf{Z}_1 = la matriz incidente para los efectos aleatorios de las repeticiones
 \mathbf{Z}_2 = la matriz incidente para los efectos aleatorios de la interacción entre repeticiones y bloques incompletos
 \mathbf{b} = el vector $i \times 1$ de los efectos aleatorios por el desempeño de los genotipos
 \mathbf{c} = vector $j \times 1$ del efecto aleatorio de las repeticiones γ_j
 \mathbf{g} = vector $j \times k$ del efectos aleatorio producto de la interacción $\gamma\delta_{jk}$
 $\boldsymbol{\epsilon}$ = vector $p \times 1$ de efectos aleatorios de los residuos.

El modelo simplificado en **R** fue: $\text{TAH} \sim (1|\text{ENTRY}) + (1|\text{REP}) + (1|\text{BLOCK:REP})$

Debido a que es un modelo lineal en donde todos los términos son considerados aleatorios, es posible encontrar un BLUP para todos los niveles de cada uno de los términos del modelo. De esta forma se obtuvieron los valores de BLUP para el efecto de las

repeticiones, el efecto de la interacción de las repeticiones con el bloque incompleto, y lo que en particular se deseaba, los BLUPs para cada nivel del efecto aleatorio genotipo. Este dato en particular se constituyó en el valor genético de cada genotipo, es decir, en su Valor de Mejora Genética (*breeding value*).

El ajuste linear del modelo se realizó mediante la función **lmer** y los valores de BLUP fueron extraídos mediante la instrucción **ranef**, ambas funciones del paquete **lme4** (Bates *et al.*, 2015), después del ajuste del modelo descrito anteriormente. Esta instrucción permite visualizar los valores de las varianzas de los efectos aleatorios de cada factor del modelo, y en el caso de “ENTRY” (Genotipo) el efecto aleatorio de cada genotipo con respecto a la variable de respuesta TAH. Este se constituyó en la mejor predicción lineal insesgada (Best Linear Unbiased Prediction) para cada genotipo con respecto a la variable TAH, o sea en un índice de desempeño, tal como lo plantean Henderson *et al.*, (1959), Robinson, (1991) y Thompson, (2008).

Para la presentación de resultados, se empleó el criterio de selección del 10% de la fracción superior de los genotipos según el valor de los índices Smith y KN. El programa **Rindsel** ofrece tres opciones 5, 10 y 20%, aunque este valor es siempre decisión del fitomejorador. Lo más común es el empleo de la selección del 10% de la fracción superior según la variable de evaluación, en este caso, índices (Pérez-Elizalde *et al.*, 2014; Pesek y Baker, 1970).

3.6.7.8 Base De Comparación Entre Índices: Competencia Del Fitomejorador

Para realizar las comparaciones necesarias, y establecer la calidad de valoración genética de los índices, se estableció lo que se denominará *Base de Comparación*. Esta base de comparación es aquella que se relaciona con la competencia del fitomejorador para realizar una selección, según su experiencia y el dominio de la actividad de mejora genética en caña de azúcar. Esta base de comparación está altamente relacionada con la lista de requisitos de los clientes para las variedades de caña de azúcar de CENGICANA (Tabla 3.3).

En general, la competencia del fitomejorador para hacer una selección apropiada y ajustada a los requerimientos del cliente, así como de acuerdo con su experiencia y conocimiento, deben cumplir, en alta proporción, con estándares como los siguientes:

Los genotipos a ser seleccionados deben:

- Estar libres o ser tolerantes a las principales enfermedades de la caña de azúcar. Este tema no se incluyó como parte de este trabajo; sin embargo, todas las evaluaciones y consideraciones necesarias se cumplieron. Las principales enfermedades de la caña de azúcar se presentan en el listado de requisitos del cliente en la Tabla 3.3.
- Presentar valores de TAH que superen y, en el peor de los casos, igualen a los que presenten los testigos propuestos, particularmente la variedad CP72-2086.
- Presentar valores de corcho y oquedad menores o al menos igual al que presenta la variedad CP72-2086.
- Presentar valores de fibra (materia seca) superiores al 12%.
- Las variedades deben presentar, a ser posible, bajas tasas de floración, inferiores al 20%.
- Presentar un equilibrio razonable entre la producción de biomasa y la concentración de azúcar. Los genotipos con alta producción de biomasa y bajas concentraciones de sacarosa no son considerados como buenos genotipos, y viceversa.
- Presentar buenos niveles de población: los valores por debajo de 12 o 13 tallos por metro lineal son considerados como malos.
- La altura de corte debe superar los 2 m. Esta variable es crucial para la acumulación de biomasa.
- Los valores del diámetro de los tallos deben ser los apropiados. Valores del diámetro por encima de los 30 mm están asociados con bajas poblaciones, y altos niveles de oquedad y corcho. Los valores inferiores a 20 mm serán de genotipos con bajos rendimientos de biomasa.
- Otros.

Todas estas variables conforman, en parte, el fundamento de la competencia del fitomejorador. Sin embargo, la postura del fitomejorador ante un genotipo también incluye su capacidad de intuir buenos genotipos. La interacción $G \times E$, puede no permitir una apreciación justa de los materiales; en este momento, la experiencia del fitomejorador permitirá establecer el potencial del material en otros posibles escenarios.

La base final de comparación fue la variable de selección (SI o NO) con base en la competencia del fitomejorador, la cual es una variable de distribución binomial.

3.6.7.9 Regresión Logit

Debido a que la variable de base de comparación es una variable binomial (SI o NO), para realizar los contrastes entre los índices y la base de comparación se utilizó el ajuste de modelos logísticos, en particular la transformación **Logit** (Manning, 2007; Everitt, 2014). La variable de transformación (endógena) puede salirse del rango entre 0 y 1. Para evitar esto, se utilizó un modelo de probabilidad no lineal, en donde se garantiza un resultado en la estimación que vaya de 0 a 1 (de NO a SI) (Manning, 2007). Se utilizó el ajuste de la función logística. La especificación de los modelos logísticos es como sigue:

$$Y_i = \frac{e^{\alpha + \beta_k X_{ki}}}{1 + e^{\alpha + \beta_k X_{ki}}} + \varepsilon_i \quad 3.14$$

El modelo **Logit** se puede interpretar en términos probabilísticos, es decir, sirve para medir la probabilidad de que ocurra o no el acontecimiento de estudio ($Y_i = \text{SI}$) (Medina, 2003).

Los modelos **Logit** se ajustaron mediante el uso de la función `glm{stats}` (R Core Team, 2017). Dentro de las líneas de programación se escribió la instrucción **family=binomial**, con el objeto de que el ajuste fuera para un modelo no lineal, pero que fue transformado en uno lineal mediante una escala logarítmica. La salida en R fue el ajuste de un modelo binomial del tipo **Logit**. Unas de las salidas importantes son los coeficientes del modelo y, adicionalmente, el valor de un pseudo- r^2 . Las salidas de R brindan tres opciones de Pseudo r^2 . La que se empleó, ya que se referencia como la más robusta, fue el valor de Pseudo r^2 de McFadden (Hu *et al.*, 2006).

Como el modelo **Logit** es un modelo probabilístico no lineal de regresión, puede escribirse según la Ecuación 3.15.

$$\ln \frac{p}{1-p} = z \quad 3.15$$

p = la probabilidad de que un suceso ocurra, en este caso de que ocurra la selección (SI)

z = el conjunto de variables independientes. En este caso la regresión logística es muy simple, porque el valor de z es asumido individualmente por los valores los índices de selección, es decir, el *score* para cada índice de selección calculado. Así pues, el modelo de regresión logística es relativamente simple: *La probabilidad de que ocurra o no el suceso de selección (SI), en escala logarítmica, es función del índice de desempeño (o de selección).* Se realizaron ajustes individuales para cada uno de los índices de selección, en cada uno de los conjuntos de datos.





4. RESULTADOS



4 RESULTADOS

4.1 RESULTADOS GENERALES

De los cuatro experimentos realizados fue posible recabar información fenotípica durante dos años (dos ciclos de cultivo y cosecha), excepto en uno de estos experimentos (**litoral_no_flor_1ª_soca**) del cual no se presentan resultados del segundo año de evaluación. La pérdida de datos en este experimento para este ciclo se debió a un incendio provocado, lo que no permitió recoger los datos correspondientes. En consecuencia, se analizaron los resultados de siete conjuntos de datos: cuatro localidades en dos años, menos la excepción de la que ya se hizo referencia (epígrafe 3.1 en Materiales y Métodos).

Para el conjunto total de datos (499 muestras) (Tabla 4.1), los °Brix (**brixprom_1**) variaron entre un mínimo de 16.7 y un máximo de 24.9, en tanto que la variable de la acumulación de la biomasa (**peso5stalk_2**) tuvo un rango que varió entre 2.9 y 12.95 kg. Los valores para las tres mediciones de “oquedad” fueron más amplios, aunque los coeficientes de variabilidad no fueron tan altos como el de **peso5stalk_2** (23.7). Las variables de medida en el laboratorio de la concentración de azúcar (**pol_6**, **brixjugo_7** y **purity_8**) presentaron valores relativamente bajos de variabilidad (CV = 9.5 y 3.3 respectivamente), en tanto que **brixprom_1** (sólidos solubles medidos en el campo) presentó una variabilidad menor a sus pares del laboratorio (CV= 7.4 y 9.5, respectivamente). Los valores más altos de variabilidad los presentaron las variables población por metro lineal (**pop_9**) (CV=21.2) y la proporción de presencia de flores (**flower_12**) (CV=26.6), esta última junto a la presencia de corcho (**pith_13**) presentaron los rangos más altos (0 a 100). Es oportuno mencionar que los coeficientes de variabilidad más altos parecen estar relacionados con variables de variación cuantitativa, de donde es preciso destacar **peso5stalk_2** (CV=23.7), **pop_9** (CV=21.2) y **pith_13** (CV=20.7).

Tabla 4.1. Medidas generales descriptivas para las variables en estudio.

Variable	Unidades de medida	Mín	Máx	Med	Media	ES	DE	CV
brixprom_1	°Brix	16.67	24.93	20.80	20.76	0.07	1.53	7.4
peso5stalk_2	kg	2.90	12.95	7.20	7.25	0.08	1.72	23.7
oqueleve_3	%	0.00	86.50	79.90	76.51	0.54	12.12	15.8
oqueinter_4	%	0.00	58.80	58.80	53.16	0.46	10.18	19.1
oqueseve_5	%	0.00	79.30	79.30	76.59	0.37	8.23	10.7
pol_6	%	7.83	18.30	15.08	14.90	0.08	1.73	11.6
brixjugo_7	°Brix	12.82	23.94	20.20	20.02	0.09	1.90	9.5
purity_8	%	73.94	97.20	90.80	90.46	0.14	3.02	3.3
pop_9	Tallos m ⁻¹	7.00	23.00	15.00	14.48	0.14	3.07	21.2
alt_10	m	1.50	3.20	2.40	2.40	0.01	0.31	12.9
diam_11	mm	14.00	37.50	25.30	25.47	0.15	3.41	13.4
flower_12	%	0.00	100.00	100.00	89.55	1.06	23.78	26.6
pith_13	%	0.00	100.00	100.00	88.22	0.82	18.23	20.7
fiber_14	%	8.95	17.61	12.73	12.95	0.07	1.63	12.6

Min= valor mínimo; Máx=Valor máximo; Med=Mediana; ES=Error estándar; DE=Desviación Estándar; CV=Coeficiente de Variación.
brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Las medidas descriptivas de las variables para cada experimento mostraron que los experimentos del estrato litoral presentaron valores más altos que los que se observaron en el estrato medio para la variable relacionada con la biomasa, particularmente entre los experimentos **flor**, mientras que las diferencias con respecto al tipo **no_flor** fueron mucho menos acentuadas (Tabla 4.2). Por otro lado, los valores de las variables relacionadas con el azúcar estuvieron a favor de los genotipos del estrato medio. En este sentido, los conjuntos de datos del experimento de variedades **no_flor** del estrato medio, tanto en plantía como en primera soca, fueron los que sumaron un mayor número de genotipos en evaluación.

Tabla 4.2. Medidas descriptivas de 15 variables por conjunto de datos. Para cada genotipo estudiado se utilizaron muestras de diez tallos.

Conjunto de datos	Número de genotipos	Variable	Unidades de medida	Media	Mediana	Desviación estándar	Máximo	Mínimo	CV
litoral_flor_plantía	40	brixprom_1	°Brix	19.38	19.40	1.03	22.70	17.00	5.31
		peso5stalk_2	kg	7.37	7.25	1.47	11.40	5.10	19.95
		oqueleve_3	%	77.63	80.25	11.97	86.50	6.30	15.42
		oqueinter_4	%	54.89	58.80	7.44	58.80	26.00	13.55
		oqueseve_5	%	76.55	79.30	7.02	79.30	44.90	9.17
		pol_6	%	13.40	13.62	1.54	15.72	9.80	11.49
		brixjugo_7	°Brix	18.61	19.03	1.97	22.32	12.82	10.59
		purity_8	%	87.45	88.50	3.95	93.13	73.94	4.52
		pop_9	Tallos m ⁻¹	18.15	18.00	2.05	23.00	14.00	11.29
		alt_10	m	2.59	2.60	0.29	3.10	1.88	11.20
		diam_11	mm	26.66	26.60	3.21	35.00	19.50	12.04
		flower_12	%	60.00	85.95	42.59	100.00	0.00	70.98
		pith_13	%	78.20	87.00	25.53	100.00	0.00	32.65
		fiber_14	%	12.74	12.61	1.79	16.27	9.33	14.05
		TAH	t ha ⁻¹	23.77	22.73	5.72	39.59	14.13	24.06
litoral_flor_1ª_soc a	31	brixprom_1	°Brix	19.62	19.70	1.15	22.20	17.20	5.86
		peso5stalk_2	kg	8.03	7.78	1.68	12.95	3.80	20.92
		oqueleve_3	%	78.29	82.40	9.68	86.50	46.40	12.36
		oqueinter_4	%	53.25	58.80	10.61	58.80	4.50	19.92
		oqueseve_5	%	74.65	79.30	10.05	79.30	25.00	13.46
		pol_6	%	13.45	13.87	1.85	16.60	7.83	13.75
		brixjugo_7	°Brix	18.32	18.60	1.57	21.20	14.30	8.57
		purity_8	%	89.14	89.75	3.08	94.33	78.83	3.46
		pop_9	Tallos m ⁻¹	15.31	16.00	2.12	20.00	9.00	13.85
		alt_10	m	2.34	2.30	0.27	3.20	1.50	11.54
		diam_11	mm	25.43	25.35	3.09	37.50	20.00	12.15
		flower_12	%	90.18	99.15	20.32	100.00	12.00	22.53
		pith_13	%	89.15	96.50	13.83	100.00	52.00	15.51
		fiber_14	%	13.37	13.39	1.44	17.41	11.06	10.77
		TAH	t ha ⁻¹	22.08	21.33	6.39	39.22	9.07	28.94
litoral_no_flor_plantía	42	brixprom_1	°Brix	21.43	21.45	1.36	24.30	18.00	6.35
		peso5stalk_2	kg	7.56	7.30	1.67	12.25	4.80	22.09
		oqueleve_3	%	77.22	80.65	12.01	86.50	26.40	15.55
		oqueinter_4	%	56.26	58.80	5.60	58.80	34.50	9.95
		oqueseve_5	%	78.55	79.30	4.30	79.30	42.30	5.47
		pol_6	%	15.43	15.55	1.39	18.20	12.00	9.01
		brixjugo_7	°Brix	20.69	20.78	1.58	23.94	16.56	7.64
		purity_8	%	90.49	90.56	2.69	94.75	81.14	2.97
		pop_9	Tallos m ⁻¹	14.72	15.00	2.43	19.00	7.00	16.51
		alt_10	m	2.50	2.50	0.26	3.20	1.90	10.40

Conjunto de datos	Número de genotipos	Variable	Unidades de medida	Media	Mediana	Desviación estándar	Máximo	Mínimo	CV
		diam_11	mm	24.71	24.30	3.25	33.00	18.00	13.15
		flower_12	%	99.51	100.00	2.17	100.00	90.00	2.18
		pith_13	%	91.70	100.00	16.49	100.00	35.00	17.98
		fiber_14	%	12.24	12.14	1.41	15.77	8.95	11.52
		TAH	t ha ⁻¹	22.85	22.05	6.50	37.65	10.61	28.45
medio_flor_plantía	30	brixprom_1	°Brix	21.16	21.00	1.07	23.50	18.20	5.06
		peso5stalk_2	kg	7.75	7.55	1.53	11.75	4.35	19.74
		oqueleve_3	%	71.93	75.90	17.24	86.50	0.00	23.97
		oqueinter_4	%	47.77	55.00	14.08	58.80	7.80	29.47
		oqueseve_5	%	72.70	79.30	12.67	79.30	18.70	17.43
		pol_6	%	15.64	15.70	1.12	17.80	13.00	7.16
		brixjugo_7	°Brix	21.00	20.98	1.14	22.98	18.12	5.43
		purity_8	%	90.46	90.68	2.15	94.13	84.80	2.38
		pop_9	Tallos m-1	11.69	11.00	2.34	17.00	7.00	20.02
		alt_10	m	2.42	2.40	0.25	3.00	1.60	10.33
		diam_11	mm	26.64	27.00	2.84	32.70	21.30	10.66
		flower_12	%	80.19	89.00	26.20	100.00	0.00	32.67
		pith_13	%	73.15	71.00	22.56	100.00	15.00	30.84
		fiber_14	%	12.96	12.75	1.63	16.86	10.24	12.58
		TAH	t ha ⁻¹	18.76	18.21	5.71	31.79	9.47	30.44
medio_flor_1a_soca	30	brixprom_1	°Brix	19.86	20.00	1.07	21.70	17.00	5.39
		peso5stalk_2	kg	5.22	5.20	1.13	8.55	2.90	21.65
		oqueleve_3	%	74.66	78.24	13.01	86.50	40.20	17.43
		oqueinter_4	%	52.51	55.80	11.39	58.80	0.00	21.69
		oqueseve_5	%	75.77	79.30	11.93	79.30	0.00	15.75
		pol_6	%	13.95	14.30	1.14	16.60	11.30	8.17
		brixjugo_7	°Brix	18.70	19.00	1.35	22.10	15.58	7.22
		purity_8	%	91.30	91.81	2.04	95.18	86.14	2.23
		pop_9	Tallos m-1	16.30	16.00	1.43	20.00	12.00	8.77
		alt_10	m	1.93	1.90	0.18	2.50	1.50	9.33
		diam_11	mm	22.06	22.00	2.93	28.70	14.00	13.28
		flower_12	%	90.53	95.00	11.62	100.00	52.00	12.84
		pith_13	%	96.45	100.00	6.59	100.00	72.00	6.83
		fiber_14	%	14.55	14.65	1.54	17.61	11.55	10.58
		TAH	t ha ⁻¹	15.71	15.31	3.65	24.66	8.82	23.23
medio_no_flor_1a_soca	49	brixprom_1	°Brix	21.82	22.00	1.30	24.93	16.67	5.96
		peso5stalk_2	kg	7.32	7.22	1.69	11.85	3.75	23.09
		oqueleve_3	%	76.49	78.80	9.92	86.50	48.60	12.97
		oqueinter_4	%	51.49	58.35	10.96	58.80	8.30	21.29
		oqueseve_5	%	78.18	79.30	3.46	79.30	57.90	4.43
		pol_6	%	16.27	16.40	0.87	18.30	14.30	5.35
		brixjugo_7	°Brix	21.63	21.59	0.92	23.64	19.54	4.25

Conjunto de datos	Número de genotipos	Variable	Unidades de medida	Media	Mediana	Desviación estándar	Máximo	Mínimo	CV
		purity_8	%	91.57	91.72	1.73	97.20	87.58	1.89
		pop_9	Tallos m-1	11.89	12.00	2.10	17.00	7.00	17.66
		alt_10	m	2.50	2.50	0.24	3.20	1.90	9.60
		diam_11	mm	25.59	25.30	2.95	32.00	20.30	11.53
		flower_12	%	98.74	100.00	8.23	100.00	38.00	8.34
		pith_13	%	94.88	100.00	13.39	100.00	12.00	14.11
		fiber_14	%	12.57	12.40	1.35	15.84	9.71	10.74
		TAH	t ha ⁻¹	18.66	18.48	4.86	34.62	9.05	26.05
medio_no_flor_plant ia	29	brixprom_1	°Brix	21.64	21.55	1.09	24.30	19.70	5.04
		peso5stalk_2	kg	7.31	7.25	1.35	11.30	4.70	18.47
		oqueleve_3	%	78.43	84.00	10.71	86.50	49.30	13.66
		oqueinter_4	%	55.76	58.80	6.75	58.80	31.80	12.11
		oqueseve_5	%	78.66	79.30	2.40	79.30	64.00	3.05
		pol_6	%	15.72	15.80	1.09	17.80	11.20	6.93
		brixjugo_7	°Brix	20.61	20.72	1.18	22.80	16.14	5.73
		purity_8	%	92.70	93.04	2.33	96.15	84.76	2.51
		pop_9	Tallos m-1	13.55	14.00	2.13	17.00	9.00	15.72
		alt_10	M	2.44	2.50	0.20	2.90	2.00	8.20
		diam_11	mm	27.50	27.70	3.22	35.70	19.30	11.71
		flower_12	%	99.12	100.00	6.70	100.00	49.00	6.76
		pith_13	%	88.33	95.50	13.32	100.00	52.00	15.08
		fiber_14	%	12.57	12.38	1.30	15.34	10.38	10.34
		TAH	t ha ⁻¹	20.63	19.95	5.19	32.73	11.50	25.16

CV = Coeficiente de Variación
 brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

4.2 BIOMASA

Para la variable Peso de una muestra de cinco tallos de caña (**peso5stalk_2**) se encontraron más de 160 genotipos que alcanzaron y superaron los 7 kg (Figura 4.1.a). Los valores de concentración de sacarosa también fueron importantes, y los principales valores de medición de esta variable fueron la refractometría en pre-cosecha (**brixprom_1**), la refractometría en el laboratorio (**brixjugo_7**) y la concentración aparente de sacarosa (**pol_6**) (Figura 4.1.b). En total, se contabilizaron poco más de 330 genotipos con concentraciones superiores a 15% de **pol_6**.

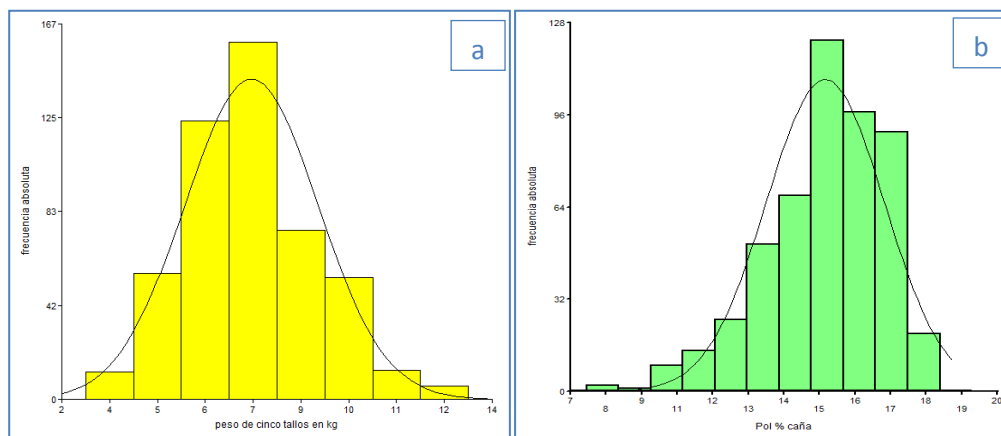


Figura 4.1. a) Número de genotipos (frecuencias) según la variable peso5stalk_2. b) Número de genotipos según el contenido aparente de sacarosa (pol_6).

El mayor número de genotipos de caña de azúcar con los mejores pesos de 5 tallos, con un promedio de 7.4 kg, se presentaron en el grupo de variedades **no_flor**, en tanto que para los genotipos **flor** el mayor número de genotipos se encontró para los valores de peso promedio de 7.12 kg; esta diferencia no fue significativa (Tabla 4.3).

Tabla 4.3. Pruebas de *t* para dos poblaciones independientes y para quince variables, agrupadas según patrón de floración (flor y no_flor) y estrato de evaluación (litoral y medio).

	Variable	Unidades de medida	Promedio flor	Promedio no_flor	Diferencia absoluta	<i>t</i>	GL	Significación
Floración	brixprom_1	°Brix	19.95	21.65	1.70	-14.88	497	****
	peso5stalk_2	kg	7.12	7.40	0.28	-1.854	497	NS
	oqueleve_3	%	75.86	77.21	1.35	-1.26	494	NS
	oqueinter_4	%	52.23	54.17	1.94	-2.16	485	*
	oqueseve_5	%	74.92	78.42	3.50	-5.04	323	****
	pol_6	%	14.04	15.84	1.80	-13.84	466	****
	brixjugo_7	°Brix	19.08	21.05	1.97	-13.7	472	****
	purity_8	%	89.53	91.47	1.94	-7.66	478	****
	pop_9	Tallos m ⁻¹	15.59	13.27	2.32	9.19	493	****
	alt_10	m	2.32	2.48	0.16	-5.99	466	****
	diam_11	mm	25.22	25.75	0.53	-1.75	497	NS
	flower_12	%	80.83	99.10	18.27	-9.67	285	****
	pith_13	%	84.61	92.19	7.58	-4.8	474	****
	fiber_14	%	13.4	12.46	0.95	6.83	489	****
	TAH_15	t ha ⁻¹	20.38	20.58	0.2	-0.37	497	NS
	Variable		Promedio estrato litoral	Promedio estrato medio	Diferencia absoluta	<i>T</i>	GL	
Estrato	brixprom_1	°Brix	20.2	21.21	1.02	-7.83	497	****
	peso5stalk_2	kg	7.62	6.95	0.67	4.42	497	****
	oqueleve_3	%	77.72	75.52	2.19	2.05	496	*
	oqueinter_4	%	54.81	51.81	2.99	3.41	492	***
	oqueseve_5	%	76.6	76.58	0.02	0.03	497	NS
	pol_6	%	14.16	15.51	1.35	-9.17	399	****
	brixjugo_7	°Brix	19.26	20.64	1.37	-8.38	417	****
	purity_8	%	89.17	91.51	2.35	-8.92	359	****
	pop_9	Tallos m ⁻¹	16.11	13.16	2.95	12.13	497	****
	alt_10	m	2.47	2.34	0.12	4.52	497	****
	diam_11	mm	25.51	25.45	0.06	0.19	497	NS
	flower_12	%	85.24	93.05	7.81	-3.49	327	***
	pith_13	%	87.05	89.18	2.12	-1.29	497	NS
	fiber_14	%	12.78	13.09	0.30	-2.09	497	*
	TAH_15	t ha ⁻¹	22.96	18.45	4.51	8.71	433	****

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo
t, Pruebas de *t* de Student; GL, grados de libertad

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH_15, t de azúcar por ha

Al realizar la comparación de la variable peso de 5 tallos (**peso5stalk_2**) entre los genotipos, según el estrato de evaluación (Figura 4.2), se encontró que en el estrato litoral se presentaron frecuencias (número de genotipos) superiores a diez, con pesos superiores a 7 kg. En tanto que en el estrato medio las frecuencias de genotipos con pesos superiores a los 7 kg estuvieron por debajo de frecuencias de diez. En este sentido, la diferencia global de 0.67 kg resultó en una diferencia altamente significativa ($P < 0.0001$).

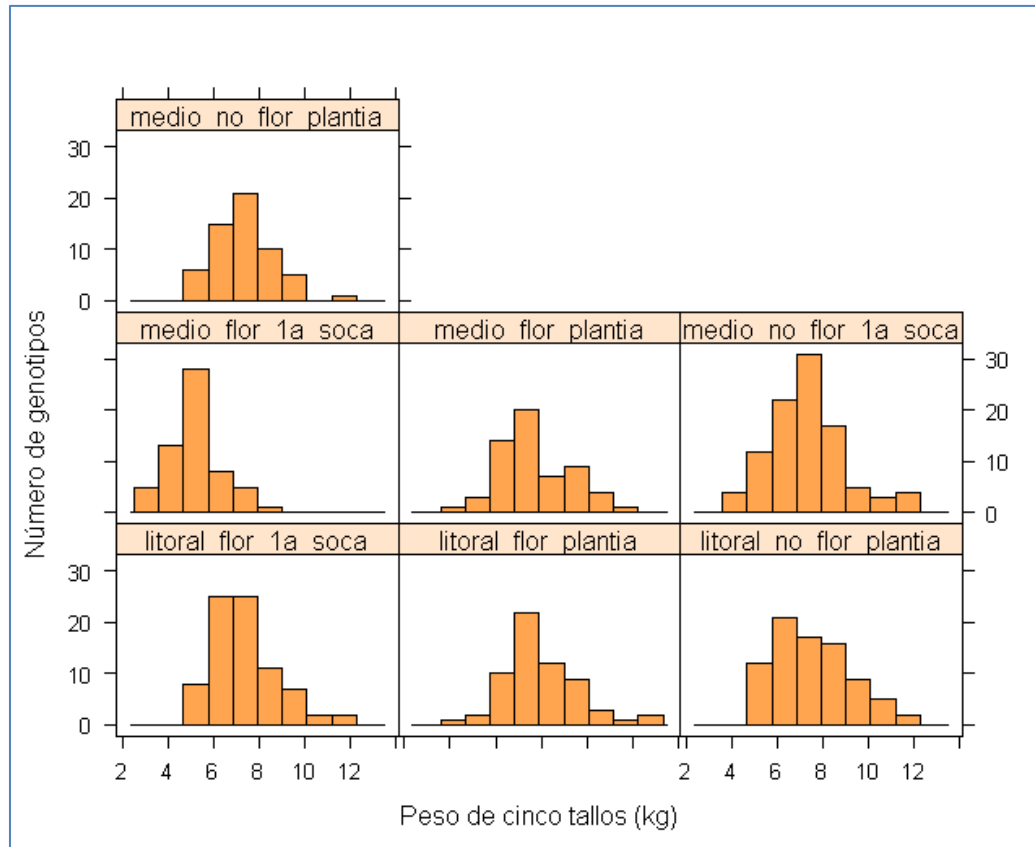


Figura 4.2 Número de genotipos para la variable peso de 5 tallos (peso5stalk_2).

Con respecto a la variable fibra (**fiber_14**), la diferencia de los promedios de los genotipos (0.30 t), cuando se compararon por estratos, fue significativa al 1% (*). En tanto que, al realizar la comparación de desempeño de los genotipos con relación al patrón de floración, esta diferencia fue altamente significativa, 0.95 t ($p < 0.0001$).

En la Figura 4.3 se observan las tendencias de mayor producción de fibra en los ensayos con los genotipos **flor** y estrato medio

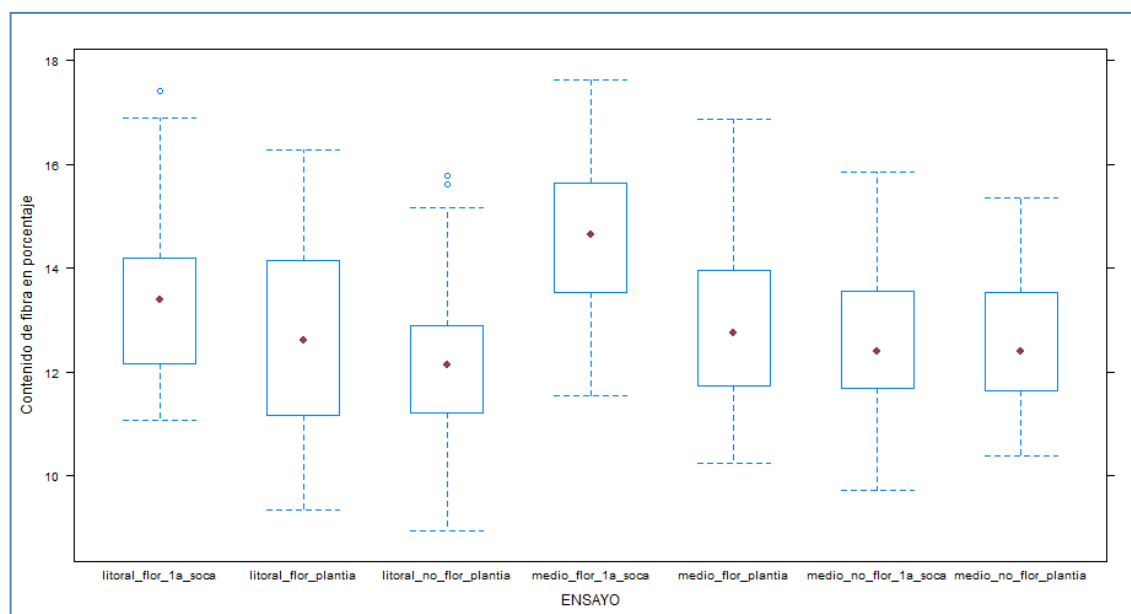


Figura 4.3. Contenido de fibra (%) (fiber_14), por conjunto de datos.

4.3 CONCENTRACIÓN DE SACAROSA

El promedio de la concentración de sacarosa (**pol_6**, Tabla 4.3 y Figura 4.4) de los genotipos del estrato medio fue significativamente mayor comparado con el promedio de los valores que presentaron los genotipos en el estrato litoral ($|t| = 9.17$, $p < 0.001$). En el estrato medio, los promedios por conjunto de datos superaron el 15% de concentración de azúcar, excepto para el conjunto **medio_flor_1^a_soca**, el cual presentó un promedio por debajo del 14%. Por otro lado, los registros de los promedios de los valores de los conjuntos de datos del estrato litoral fueron inferiores al 14%, excepto para el conjunto de datos **litoral_no_flor_plantía**, el cual presentó un valor promedio por encima del 15%.

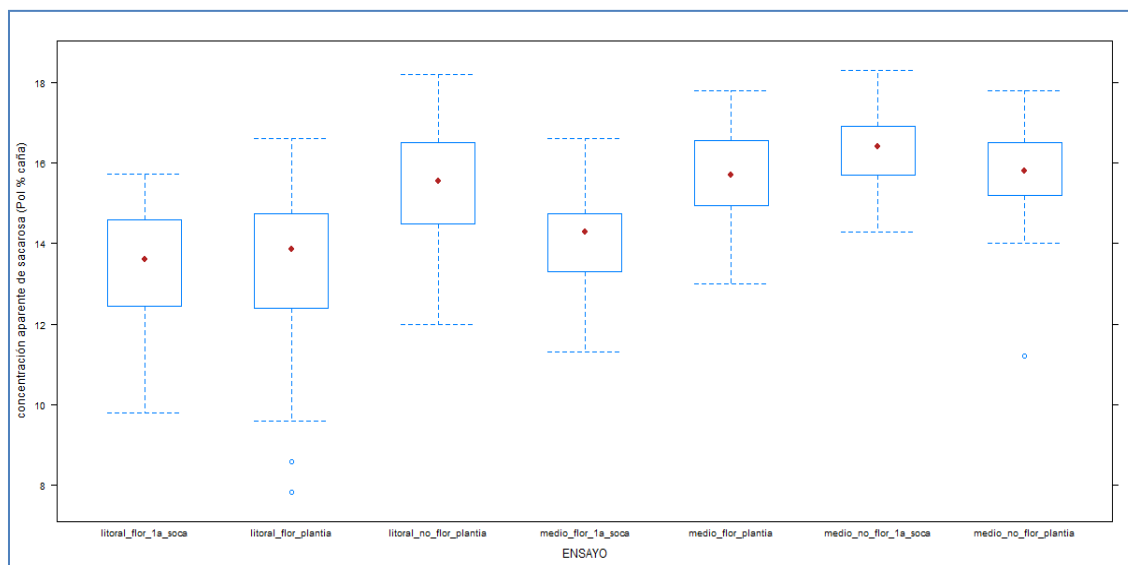


Figura 4.4. Promedio de valores del pol % caña (pol_6) por conjunto de datos.

Las mejores concentraciones de azúcar se alcanzaron en el estrato medio (Tabla 4.3, Figura 4.5.) en términos de **brixprom_1**. Las diferencias con respecto a los valores del estrato litoral fueron significativas ($|t|=7.83$; ***, Tabla 4.3).

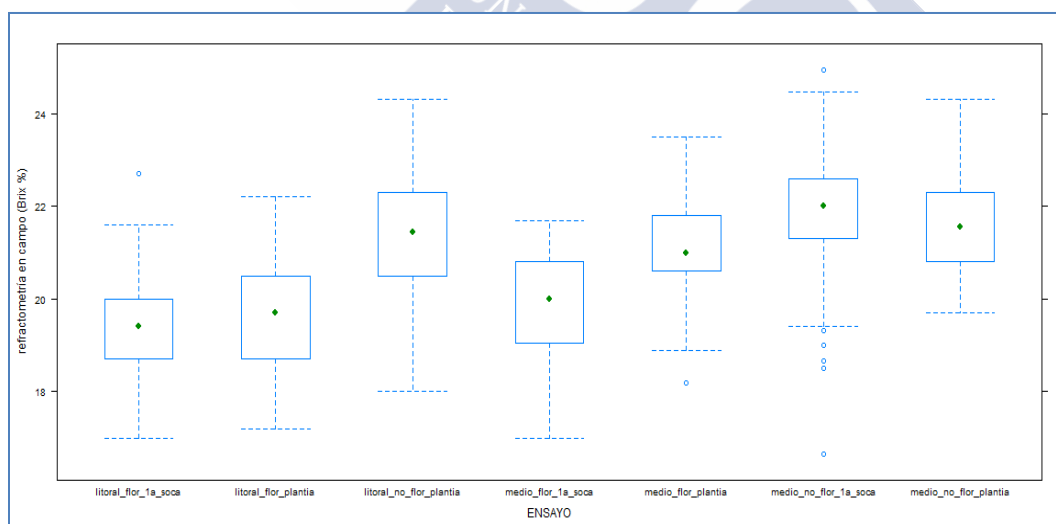


Figura 4.5. Promedios °brix en % medidos en el campo (brixprom_1).

La medición de sólidos solubles en porcentaje (**brixprom_1** y **brixjugo_7**) mostraron que el ensayo con frecuencias de hasta 50 genotipos con valores de 21% se presentaron en el

experimento **medio_no_flor_1ª_soca** (Figura 4.6), en concordancia con los resultados obtenidos para las variables Pol % caña (**pol_6**) y °Brix medido en el campo (**brixprom_1**).

Se presentaron frecuencias importantes de genotipos con altos valores de refractometría medida en el laboratorio (Figura 4.6). Los conjuntos de datos con frecuencias iguales o menores a 20 fueron: **medio_no_flor_plantía**, **litoral_flor_plantía** y **medio_flor_plantía**. En tanto que los conjuntos de datos en donde se llegaron a registrar frecuencias entre 20 y 30 para valores de **brixjugo_7** de entre 18 y 20 °Brix fueron **litoral_flor_1ª_soca**, **litoral_no_flor_plantía** y **medio_flor_1ª_soca**. En términos generales, se encontraron diferencias altamente significativas (1.37, $p < 0.0001$) con respecto a los promedios entre los grupos formados por estratos (Tabla 4.3), siendo el conjunto de datos **medio_no_flor_1ª_soca** el que más influyó para que estas diferencias fueran finalmente significativas (Figura 4.6). Por otro lado, las comparaciones entre genotipos según su patrón de floración mostraron que la diferencia, 1.97, fue también altamente significativa ($p < 0.0001$) (Tabla 4.3).

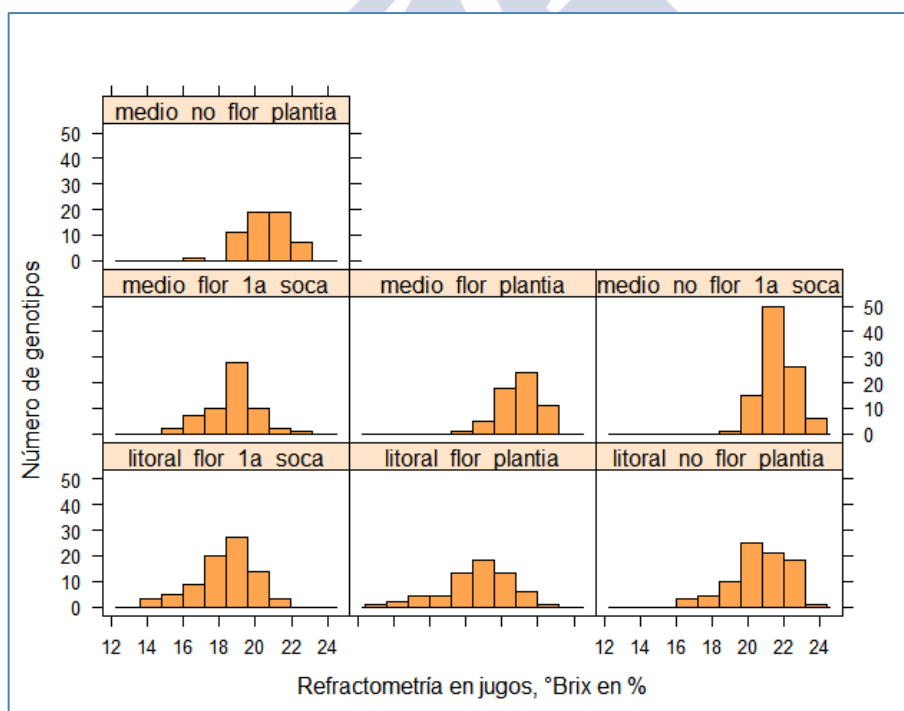


Figura 4.6. Número de genotipos según °brix medidos en el laboratorio de jugos (brixjugo_7).

4.4 ANÁLISIS DE VARIANZA

En la Tabla 4.4 se presenta el resumen de todos los análisis de varianza combinados realizados para las 14 variables de valoración y el TAH para un modelo lineal en donde todos los factores se consideraron fijos (ver modelo 3.2 en la sección de Materiales y Métodos). Los ajustes de los modelos, según la columna que indica los valores de r^2 , fueron de intermedios a altos. Para todas las variables incluidas en todos los análisis se obtuvo que, para la fuente de variación “Variedad” (ENTRY), las diferencias en los promedios fueron altamente significativas, por lo que se sigue que las diferencias en los promedios de todas las variables estudiadas fueron funciones de los materiales genéticos.

Por otra parte, para la fuente de variación “Ciclo” (Tabla 4.4), es decir el año de evaluación, también se encontraron diferencias significativas (****) para todas las variables estudiadas, excepto para las variables **oqueinter_4** y **purity_8**, en las cuales las diferencias fueron al 5% (*) así como **oqueseve_5** y **oqueleve_3**, las cuales no presentaron diferencias significativas y. Lo anterior también puede ser atribuido a un efecto puramente genético, puesto que los genotipos evaluados para cada ciclo en cada experimento fueron diferentes, excepto para los cultivares “testigo”, los cuales generaron el error experimental necesario para la elaboración de los análisis de variación combinados.

Para las diferencias en términos de las localidades (“Estrato”) se registraron diferencias estadísticas en los promedios de todas las variables, excepto para **oqueseve_5**, **diam_11**, **pith_13** y **fiber_14**, y lo cual fue indicativo de que estas variables no estuvieron afectadas en forma importante por el ambiente..

Tabla 4.4. Resumen de los análisis de variación para las variables en estudio.

Fuente de Variación en el modelo										
Variable	Estadístico	Ciclo	Estrato	Floración	Repetición	Variedad	Ciclo * estrato	Ciclo * Floración	Repetición * Bloque incompleto	r ²
brixprom_1	F valor	35.96	205.00	316.31	0.03	3.09	0.11	3.91	1.31	0.66
	Significación	****	****	****	NS	****	NS	NS	NS	
peso5stalk_2	F valor	71.2	24.16	9.06	2.21	3.73	0.55	41.57	0.70	0.57
	Significación	****	****	****	NS	****	NS	****	NS	
oqueleve_3	F valor	0.08	6.81	5.13	0.84	2.30	0.65	0	1.25	0.35
	Significación	NS	*	*	NS	****	NS	NS	NS	
oqueinter_4	F valor	5.45	16.22	13.43	0.01	3.04	0.36	3.03	0.69	0.46
	Significación	*	****	****	NS	****	NS	NS	NS	
oqueseve_5	F valor	0.32	0.01	26.49	0.91	1.25	0.52	0	0.79	0.12
	Significación	NS	NS	****	NS	NS	NS	NS	NS	
pol_6	F valor	16.22	282.65	286.48	0.37	3.47	1.08	20.37	1.31	0.69
	Significación	****	****	****	NS	****	NS	****	NS	
brixjugo_7	F valor	23.08	237.28	274.59	0.01	3.29	0.03	43.93	1.2	0.67
	Significación	****	****	****	NS	****	NS	****	NS	
purity_8	F valor	6.22	177.96	77.41	3.7	3.32	7.91	9.29	1.56	0.60
	Significación	*	****	****	NS	****	*	****	NS	
pop_9	F valor	50.25	379.8	76.07	0.02	2.49	0.88	54.43	0.94	0.63
	Significación	****	****	****	NS	****	NS	****	NS	
alt_10	F valor	104.38	21.21	76.29	4.46	3.01	0.08	59.14	0.6	0.56
	Significación	****	****	****	*	****	NS	****	NS	
diam_11	F valor	60.14	2.05	1.86	2.61	3.81	6	10.03	0.97	0.55
	Significación	****	NS	NS	NS	****	*	****	NS	
flower_12	F valor	28.42	16.34	147.31	0.25	2.35	19.14	9.22	1	0.49
	Significación	****	****	****	NS	****	****	****	NS	
pith_13	F valor	91.64	0	79.44	3.09	4.05	1.36	13.41	1.06	0.61
	Significación	****	NS	****	NS	****	NS	****	NS	
fiber_14	F valor	61.47	2.35	93.58	0.01	3.59	5.77	17.23	1.23	0.58
	Significación	****	NS	****	NS	****	*	****	NS	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

4.5 FLORACIÓN

Con respecto a los patrones de floración y su influencia en los resultados generales en todos los ensayos, también se encontraron diferencias importantes en todas las variables, tanto las relacionadas a la acumulación de biomasa como las relacionadas con la concentración de

azúcar (Tabla 4.4). Para todas las variables se presentaron diferencias estadísticas al 1% (**), excepto para la variable **oqueleve_3** la cual presentó diferencias al 5% (*) y la variable **diam_11**, la cual no presentó diferencias estadísticas (NS).

4.6 TAH (T DE AZÚCAR POR HA)

Una de las variables más importantes en los procesos de la mejora genética de caña de azúcar es la producción de azúcar (sacarosa) por unidad de área, en el caso de este trabajo t de azúcar por ha (TAH). En la sección de Introducción se hizo referencia a la importancia de esta variable, y en la sección de Materiales y Métodos se explicó la naturaleza de la misma y la forma de cálculo.

Como se explicó en la sección de Materiales y Métodos y en la sección de Introducción, el TAH es una función la acumulación de biomasa y la concentración de azúcar. En este sentido, y con los resultados observados hasta aquí, se sigue que los valores de TAH fueron influidos primeramente por la acumulación de biomasa y, en segundo lugar, por la concentración de sacarosa (epígrafes 4.2 y 4.3 respectivamente). Los experimentos ubicados en el estrato litoral presentaron mayores valores de producción de TAH ($\bar{x}=22.96 \text{ t ha}^{-1}$) con respecto a los que se presentaron en el estrato medio ($\bar{x}=18.45 \text{ t ha}^{-1}$) del área de producción de caña de azúcar en Guatemala (Figura 4.7); esta diferencia fue significativa ($t=8.71$, ****, Tabla 4.3)

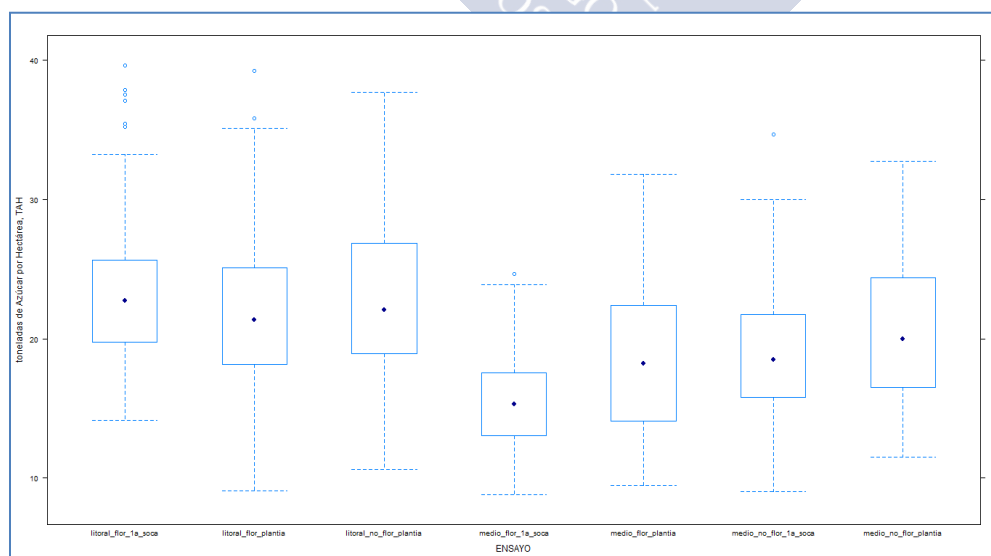


Figura 4.7. Promedios para la variable (índice) TAH para cada uno de los experimentos evaluados.

4.7 CORRELACIONES

Las correlaciones más altas se encontraron entre aquellas variables relacionadas con la concentración del azúcar (Tabla 4.5). Cada una de estas variables fue medida de forma completamente independiente una de otra a fin de reducir en el máximo posible la “colineariedad”. Las correlaciones que más destacaron fueron los valores de refractometría (°Brix), tanto en el campo (**brixprom_1**) como en el laboratorio (**brixjugo_7**), con el valor de Pol%caña (**pol_6**) (Tabla 4.5, Figura 4.8). Para el caso de las correlaciones entre **pol_6** y **brixjugo_7** se encontró el valor más alto de correlación ($r=0.95$; $Pr>F: <0.0001$), en tanto que para la correlación con **brixprom_1** el valor fue de 0.71 (***) (Tabla 4.5). Lo mismo ocurrió con la pureza del jugo (**purity_8**) versus el **pol_6** ($r=0.67$, ***). En general, se encontraron valores altos de correlación entre estas variables, lo cual es un signo general de estabilidad. Lo importante es que los valores de medición de los °Brix en el campo (**brixprom_1**) se correlacionaron positivamente con los valores de que se obtuvieron en el laboratorio (**pol_6** y **brixjugo_7**) ($r=0.95$; ***).

Tabla 4.5. Valores de correlación R de Pearson y nivel de significación

	brixprom_1	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14
brixprom_1		NS	NS	NS	*	****	****	****	****	**	NS	****	NS	**
peso5stalk_2	0.03		NS	*	NS	NS	NS	*	*	****	****	NS	NS	****
oqueleve_3	0.03	0.02		NS	**	NS	NS	NS	*	NS	NS	NS	*	NS
oqueinter_4	0.02	-0.10	0.03		****	NS	NS	NS	NS	NS	NS	NS	NS	NS
oqueseve_5	0.10	-0.08	-0.14	0.31		***	**	**	NS	*	***	NS	NS	NS
pol_6	0.68	0.01	0.01	-0.01	0.15		****	****	****	NS	NS	**	*	NS
brixjugo_7	0.67	0.03	0.01	-0.02	0.13	0.95		****	****	*	NS	*	NS	*
purity_8	0.46	-0.09	-0.01	0.02	0.15	0.74	0.58		****	**	NS	***	****	*
pop_9	-0.40	-0.11	0.10	0.09	0.00	-0.45	-0.47	-0.18		*	****	NS	NS	****
alt_10	0.14	0.49	0.08	0.07	0.09	0.07	0.11	-0.14	-0.11		****	NS	****	***
diam_11	0.03	0.58	-0.03	-0.08	-0.15	0.04	0.05	-0.04	-0.18	0.31		NS	NS	****
flower_12	0.19	-0.05	-0.06	-0.08	0.07	0.13	0.11	0.17	-0.06	-0.05	-0.03		****	****
pith_13	0.08	-0.04	-0.09	0.00	0.03	0.11	0.07	0.23	0.04	-0.18	-0.09	0.41		*
fiber_14	-0.12	-0.26	-0.01	0.05	-0.04	-0.08	-0.10	0.09	0.19	-0.17	-0.32	-0.20	-0.10	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Otras correlaciones significativas fueron las relacionadas con la acumulación de biomasa, como **peso5stalk_2** con la altura de corte (**alt_10**) ($r=0.58$; ***), así como con el diámetro medio de los tallos (**diam_11**) ($r=0.49$; ***). Un genotipo florecedor es aquel en donde el 30%, o más, de los tallos presentan flor (Orozco *et al.*, 2012), por lo que la correlación entre la producción de flores (**flower_12**) y la formación de corcho (**pith_13**) ($r=0.41$; ****) indica que la presencia de flores en más del 30% de la población en un genotipo en particular presentará formación de corcho en un 41%. Esto resulta de particular importancia, puesto que la formación de flores está correlacionada, en parte, a un carácter indeseable, como es la formación de corcho.

Las correlaciones por el método de Pearson entre el **brixjugo_7** y el **brixprom_1** ($r=0.67$, ****) y **pol_6** vs **brixprom_1** ($r=0.68$, ****), variaron en poco en los valores de r (0.01) (Tabla 4.5), considerando el total de datos. En las regresiones entre estas variables según el ciclo o el estrato de evaluación (Figura 4.8 y 4.9) se observó que las tendencias fueron similares a las que se observaron con respecto al análisis sobre el total de los datos (Tabla 4.5).

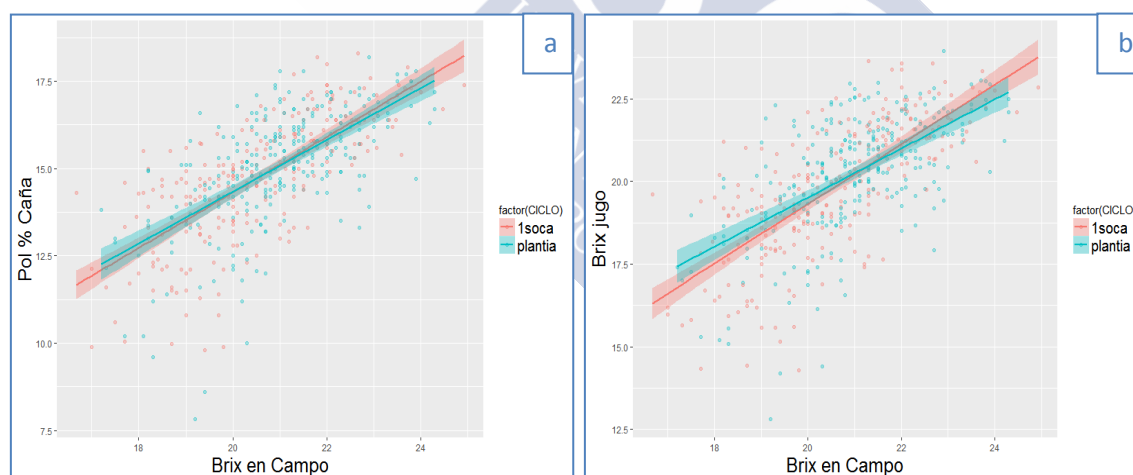


Figura 4.8. Regresión entre variables relacionadas con la concentración de azúcar: a) Pol % Caña Vs °Brix Campo (pol_6 vs brixprom_1) ($r^2= 0.4626$, ***, total de datos) y b) °Brix Jugo Vs °Brix en campo (brixjugo_7 vs brixprom_1) ($r^2= 0.4456$, ***, total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza

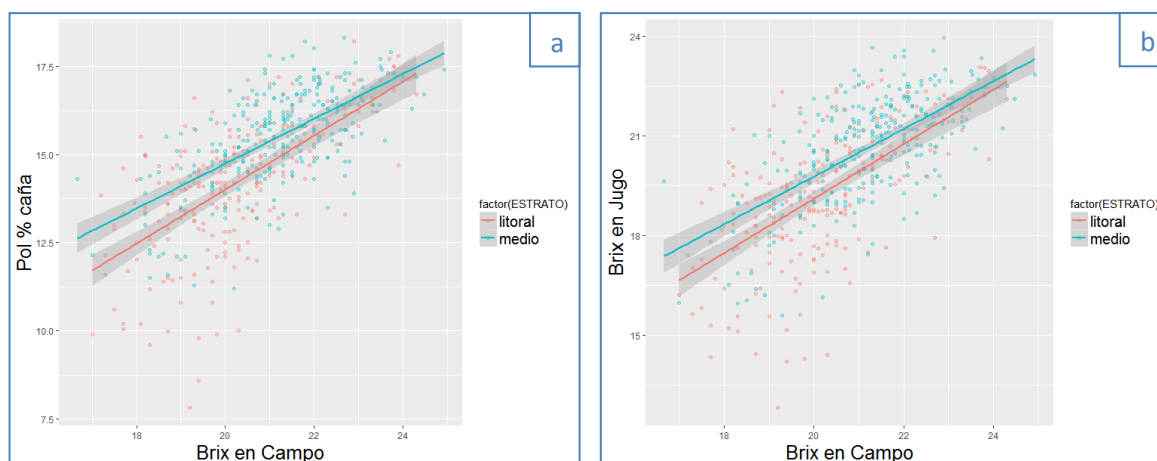


Figura 4.9. Regresión entre variables relacionadas con la concentración de azúcar: a) Pol % Caña vs °Brix campo ($r^2 = 0.476$, *** total de datos) y b) °Brix Jugo en laboratorio Vs °brix campo ($r^2 = 0.504$, *** total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza.

Con respecto a las variables relacionadas con la biomasa, éstas se correlacionaron positivamente entre sí, principalmente la variable **peso5stalk_2** con las variables **altura alt_10** y **diam_11**. Los valores de regresión entre estas variables fueron relativamente bajos, tanto para la relación **peso5stalk_2** versus **alt_10** ($r^2 = 0.24$) como para la relación **peso5stalk_2** versus el **diam_11** ($r^2 = 0.34$), y marcaron la tendencia lógica de la acumulación de biomasa. Estas relaciones no fueron muy altas y existe la posibilidad de que la regresión pueda asumir valores diferentes a los esperados según las series de datos analizadas (Figuras 4.10a y 4.10b).

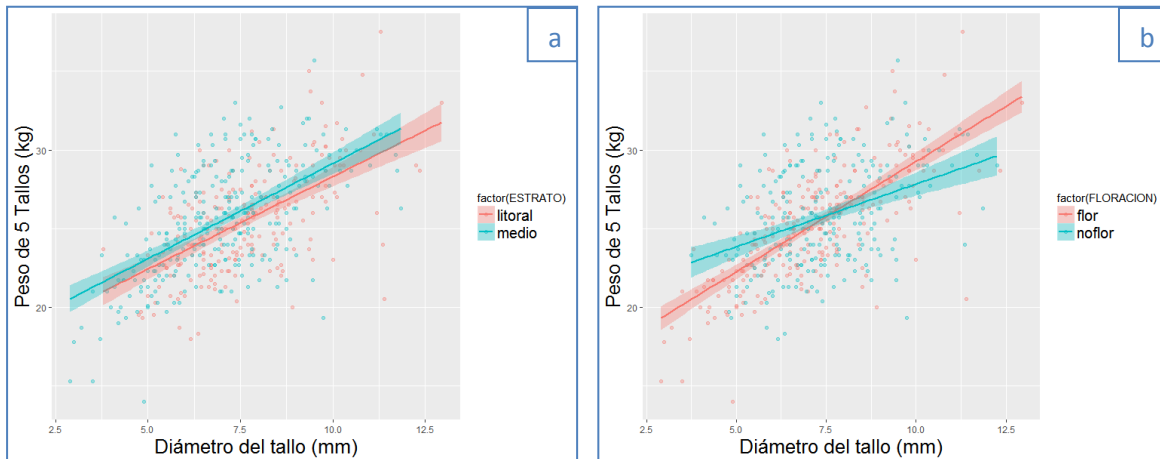


Figura 4.10. Regresión entre las variables Peso de cinco tallos vs. Diámetro según: a) el estrato de evaluación (litoral y medio) ($r^2=0.24$, ***, total de datos); y b) según el tipo de Floración (flor y no_flor) ($r^2= 0.34$, ***, total de datos). La región sombreada en cada línea de regresión indica el intervalo de confianza.

Por último, una relación extraña que se presentó continuamente en todo este trabajo, y que será discutida también en la sección de ACPs, fue la correlación negativa entre la población por metro lineal (**pop_9**) y la concentración de azúcar. Con el **pol_6** se obtuvo $r = -0.45$ (****) (Figura 4.11, Tabla 4.5). Con **brixjugo_7** el resultado fue $r = -0.47$, ****; y para **brixprom_1**, $r = -0.40$, ****.

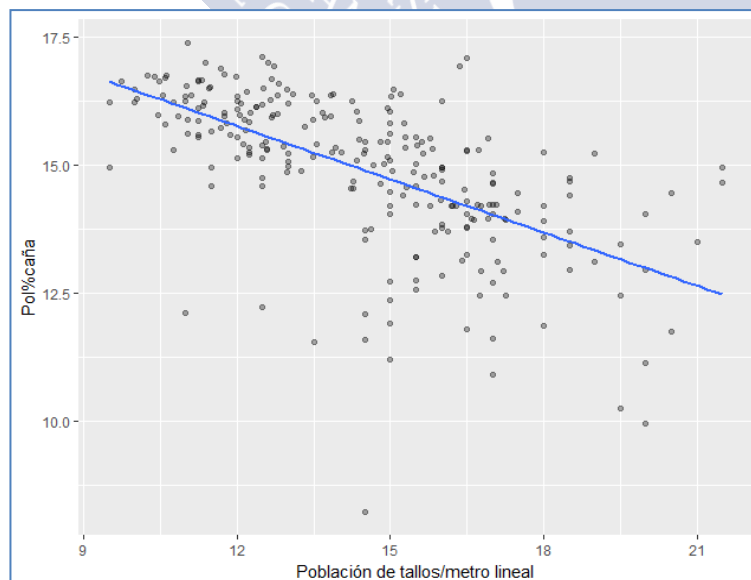


Figura 4.11. Regresión Pol % caña vs población de tallos por metro lineal ($r^2=0.202$, ***)

Por otro lado, se encontró una correlación positiva de $r = 0.41$ (****) entre las floración (**flower_12**) y corcho (**pith_13**) (Tabla 4.4). Esta correlación realmente es negativa, debido a la transformación que se hizo de la variable Corcho (**pith_13**) (ver el procedimiento en la sección de Materiales y Métodos). Se confirmó que valores altos de floración pueden eventualmente estar asociados a la generación de corcho en el tallo. Sin embargo, también es cierto que no ocurre en todos los casos, por cuanto el valor de r^2 es relativamente bajo (0.168, ***), lo que significa que la presencia de corcho está solamente parcialmente influida por el apareamiento de flores y puede considerarse, en buena medida, como una condición genética.

4.8 ANÁLISIS GENERAL COMBINADO

El análisis de variación combinado mostró, en principio, que la fuente de variación relacionada con el material genético (ENTRY) fue altamente significativa (Tabla 4.6), además de las fuentes de variación CICLO (plantía y primera soca), ESTRATO (litoral y medio), y FLORACIÓN (**flor** y **no_flor**), las cuales también mostraron diferencias significativas.

Tabla 4.6. Resumen de los resultados del Análisis de Varianza (ANOVA) combinado para la variable t de azúcar por ha (TAH).

Fuente de Variación	TAH		
	F	Prob F	Significación
CICLO	15.70	9.09^{-05}	****
ESTRATO	108.70	2.20^{-16}	****
FLORACIÓN	6.36	0.012	**
REPETICIÓN (REP)	0.85	0.357	NS
VARIEDAD (ENTRY)	2.38	1.72^{-11}	****
CICLO:ESTRATO	2.11	0.148	NS
CICLO:FLORACIÓN	1.84	0.176	NS
REPETICIÓN:BLOQUE INCOMPLETO	0.82	0.690	NS

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

Al considerar todas fuentes de variación, excepto la variable relacionada con los genotipos (ENTRY), como efectos aleatorios en un modelo lineal mixto, se encontró también un alto valor de significación estadística para la fuente genotipos (ENTRY) con efectos fijos ($F=2.436$; $\text{Prob}>F= 1.691^{-11}$). Por lo que se sigue que existe un fuerte efecto genético en el conjunto total de datos, producto del desempeño individual de cada genotipo en estudio. Esto es particularmente importante por cuanto se deduce que se cuenta con variabilidad genética

que ejerce un efecto y, además, que se cuenta con material genético para ser evaluado y sujeto de selección.

De acuerdo con los resultados obtenidos (Tabla 4.6), se encontraron dos tipos de efectos bien marcados, por un lado el efecto ambiental (ESTRATO, FLORACIÓN, REPETICIÓN) que se refleja en las diferencias entre los componentes del modelo relacionado; y, por otro lado, el efecto genético (Variedad, ENTRY) representado en los genotipos en estudio, así como en las correspondientes interacciones de estos efectos.

4.9 ESTIMACIÓN DE LOS VALORES DE HEREDABILIDAD DE LAS CARACTERÍSTICAS EN ESTUDIO

Los valores más bajos de heredabilidad se encontraron en el ensayo ubicado en el estrato medio de las variedades **no_flor** (**medio_no_flor** y **medio_no_flor**) (Tabla 4.7), lo que se correspondió con el mayor número de genotipos con nivel similar de expresión fenotípica, tanto en acumulación de biomasa (**peso5stalk_2**) como en las variables relacionadas con la concentración de azúcar (**pol_6**, **brixprom_1**) y, sobre todo, con respecto al TAH. En este sentido, la diferenciación entre genotipos fue relativamente baja, como consecuencia de una menor variación en los valores de las variables en estudio, en particular, las relacionadas con la concentración de azúcar (Tablas 4.2 y 4.3; Figuras 4.4, 4.5 y 4.6).

Las bajas diferencias entre genotipos, según las variables medidas, generaron valores más bajos en las varianzas (datos no mostrados). En consecuencia, los valores de heredabilidad también se redujeron como consecuencia de los menores valores de la Varianza Genética (σ_g^2). Al observar otros patrones se apreció que los genotipos **no_flor** presentaron valores de heredabilidad menores que los genotipos **flor** para las variables relacionadas con la concentración de azúcar. Lo anterior es consistente con el hecho de que los promedios menores de concentración de azúcar en los genotipos **flor** presentaron valores mayores de variación (Tabla 4.3), lo cual incide en un valor más alto de heredabilidad. En este sentido, estas variables parecen ser más útiles en la valoración de los genotipos **flor** que lo que podría ser en los genotipos **no_flor**.

La heredabilidad para la variable “Población” (ESTRATO) y para los genotipos florecedores (**flor**) resultó igual a cero, lo que puede estar relacionado con que las variaciones

entre genotipos con respecto a esta variable fueron relativamente bajas (Tabla 4.2 y 4.3), y a que en el grupo de genotipos evaluados no se presentaron diferencias importantes con respecto a esta variable (Tabla 4.2).

Las variables más estables en términos de sus valores de heredabilidad fueron **peso5stalk_2**, **diam_11** y **pith_13**, las cuales están relacionadas con la producción de biomasa y calidad interna de los tallos.

La variable presencia de flor (**flower_12**) presentó un valor alto en el experimento litoral **no_flor** evaluado en dos ciclos, sin embargo, debido a que esta variable está transformada (ver la sección Materiales y Métodos, epígrafe 3.5), en realidad su heredabilidad es cero ("0"). El valor de floración para todos los genotipos **no_flor** fue cero y se comparó con dos genotipos testigo con valores de floración inferiores al 10%. Es decir, que no se generó ninguna variación por cuanto todos los genotipos **no_flor** no presentaron flor en los dos años de evaluación. Por todo lo anterior, para todos los cálculos de los índices respectivos, esta variable no fue utilizada en los genotipos **no_flor**, puesto que no ofreció valores importantes de variación entre los genotipos en estudio (excluyendo a los testigos) y no resultó en un ponderador de utilidad.

Tabla 4.7. Valores estimados de heredabilidad para cada variable en estudio en tres experimentos con evaluaciones realizadas durante dos años (litoral_flor, medio_flor, medio_no_flor) y un año (litoral_no_flor).

Variable	Heredabilidad (H^2)			
	litoral_flor	litoral_no_flor	medio_flor	medio_no_flor
brixprom_1	0.70	0.82	0.72	0.00
peso5stalk_2	0.84	0.70	0.79	0.62
oqueleve_3	0.52	0.89	0.49	0.70
oqueinter_4	0.67	0.69	0.58	0.70
oqueseve_5	0.54	0.00	0.19	0.46
pol_6	0.77	0.67	0.73	0.27
brixjugo_7	0.76	0.48	0.82	0.30
purity_8	0.84	0.34	0.66	0.59
pop_9	0.00	0.66	0.00	0.69
alt_10	0.65	0.36	0.50	0.59
diam_11	0.70	0.83	0.62	0.82
flower_12	0.70	1.00	0.68	0.38
pith_13	0.72	0.86	0.59	0.90
fiber_14	0.79	0.52	0.70	0.84

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Para los conjuntos de datos del estrato medio y los genotipos **no_flor**, las varianzas relacionadas con los genotipos (ENTRY) y los ciclos de evaluación fueron muy bajas, 9.40^{-16} y $0.00E+00$, respectivamente y, por tanto, el valor de heredabilidad es prácticamente cero

4.10 ANÁLISIS DE COMPONENTES PRINCIPALES (ACP)

Ya que se hizo la medición y ajuste de 14 variables de caracterización en todos los ensayos (Materiales y Métodos, epígrafe 3.5), se realizaron una serie de análisis de estadística multidimensional con el fin de conocer la dinámica y estructura de las observaciones en función de las variables utilizadas para caracterizar a los genotipos en evaluación. Uno de estos análisis fue el de Análisis de Componentes Principales (ACP). Este análisis se realizó en principio para reducir la dimensionalidad original y, al mismo tiempo, maximizar la variabilidad (Hernández, 1998; Dunteman, 1984; Huntley, 2011). Por otra parte, este análisis sirvió para establecer la importancia de las variables en términos de su organización por grupos, y sus propiedades con respecto a los genotipos de caña de azúcar en evaluación. Además, los *scores* de los ACPs fueron empleados como funciones lineales de las variables en estudio, es decir, como índices de desempeño o de selección.

En un ACP general, donde se consideraron todos los datos de los siete conjuntos, con las primeras cuatro CPs se alcanzó a cubrir una importante proporción del total de la variabilidad identificada (60.79%) (Tabla 4.8, Figura 4.12).

Las dos primeras CPs explicaron el 40.60% (Tabla 4.8 y Figura 4.12) y las tres primeras CPs explicaron el 51.21% del total de la variabilidad (Tabla 4.8 y Figuras 4.12 y 4.15). Finalmente, las seis primeras CPs explicaron hasta el 74.55% del total de la varianza. El valor total de la varianza fue de 14.00 (Tabla 4.8).

Tabla 4.8. *Eigenvalues* (varianzas) y su participación porcentual en cada CP para todos los datos.

Componente	Varianza	Porcentaje de la varianza	Porcentaje acumulado
CP1	3.452	24.660	24.660
CP2	2.230	15.930	40.590
CP3	1.486	10.620	51.210
CP4	1.342	9.588	60.793
CP5	1.057	7.550	68.340
CP6	0.868	6.203	74.546
CP7	0.751	5.363	79.909
CP8	0.667	4.762	84.671
CP9	0.554	3.955	88.626
CP10	0.531	3.790	92.420
CP11	0.396	2.830	95.250
CP12	0.344	2.459	97.705
CP13	0.301	2.147	99.851
CP14	0.021	0.149	100.000
Total de la varianza	13.999	100.00	

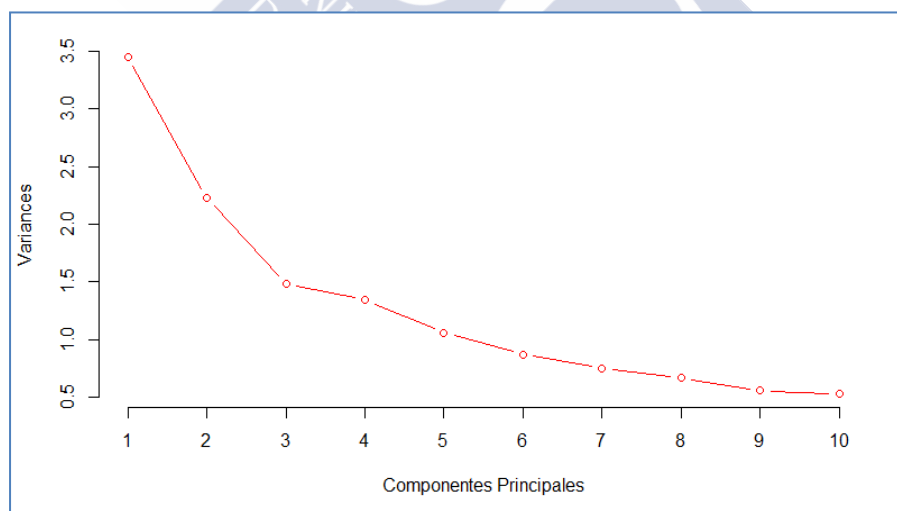


Figura 4.12. Varianzas por CP al considerar todas las observaciones de los siete conjuntos de datos.

Prácticamente todas las variables originales tuvieron una importancia en al menos una de las primeras cuatro CP (Tabla 4.9, Figura 4.13), aunque el aporte en cada componente no fue del mismo orden. Más adelante se discuten las cargas para las primeras tres CPs. La CP4 estuvo más influida con las características de llenado del tallo (oquedad). Hay que señalar que

las CP3 y CP4 explicaron, cada una, cerca del 10% del total de la variabilidad, de ahí su relativa importancia (estas tendencias también se observaron en las Figuras 4.13 y 4.14).

Tabla 4.9. Cargas (*eigenvectors* o *loadings*) de cada variable original con respecto a las primeras seis CPs.

	CP1	CP2	CP3	CP4	CP5	CP6
brixprom_1	0.430	-0.002	0.063	-0.013	0.076	-0.134
peso5stalk_2	0.046	0.543	-0.063	0.076	0.058	0.375
oqueleve_3	-0.017	0.046	0.190	-0.183	0.860	-0.198
oqueinter_4	-0.007	-0.109	0.285	0.583	0.162	0.020
oqueseve_5	0.102	-0.136	0.147	0.651	-0.136	-0.054
pol_6	0.511	-0.058	0.113	-0.071	0.021	0.082
brixjugo_7	0.492	-0.017	0.132	-0.078	-0.001	-0.016
purity_8	0.391	-0.217	0.011	-0.055	0.084	0.406
pop_9	-0.308	-0.161	-0.032	0.144	0.335	0.409
alt_10	0.069	0.439	0.190	0.265	0.141	0.047
diam_11	0.062	0.514	-0.101	-0.020	-0.069	0.278
flower_12	0.143	-0.091	-0.581	0.164	0.175	-0.140
pith_13	0.105	-0.179	-0.573	0.157	0.163	0.287
fiber_14	-0.103	-0.315	0.332	-0.197	-0.066	0.528

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Cuando se analiza la distribución de los valores (*scores*) de las dos primeras CPs para todos los conjuntos de datos (Figura 4.13), las cuales acumulan el 40.6% de la varianza total general, se distingue cómo los valores (*scores*) de las CPs diferencian a todos los siete conjuntos de datos. El conjunto de datos **medio_flor_1ª_soca** presentó los niveles más bajos de expresión fenotípica, lo que también se vio reflejado en los *scores* para las primeras dos CPs.

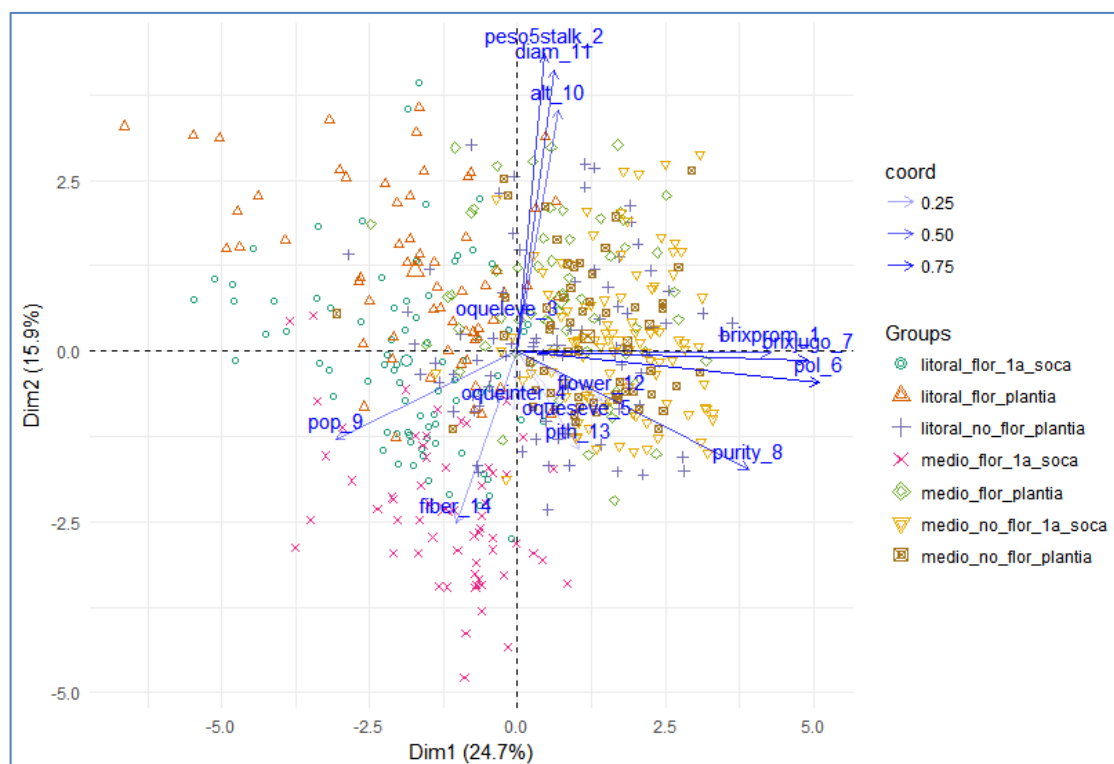


Figura 4.13. Distribución de los siete conjuntos de datos del estudio según las dos primeras CPs (Dim1 vs Dim2) que acumularon el 40.6% de la variabilidad.

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Según lo mostrado en la Tabla 4.9 y la Figura 4.13, los genotipos fueron agrupados según lo plantean los vectores correspondientes de las variables originales. Para la componente CP1, los vectores que dominaron fueron los relacionados con la concentración de azúcar (**brixprom_1, pol_6, brixjugo_7, purity_8**) (Tabla 4.9 y Figura 4.14); para la CP2, fueron las variables relacionadas con la producción de biomasa (**peso5stalk_2, diam_11, alt_10 y fiber_13**) (Tabla 4.9 y Figura 4.14). Lo anterior está relacionado con que los genotipos con altas concentraciones de azúcar fueron principalmente los genotipos **no_flor** evaluados en el estrato medio, tanto en plantía como en 1a soca (Tabla 4.2 y Tabla 4.3). En la CP3, las variables que más contribuyeron fueron las relacionadas con la calidad del tallo (**flower_12, pith_13, fiber_14 y oqueinter_4**) (Tabla 4.9 y Figura 4.14).

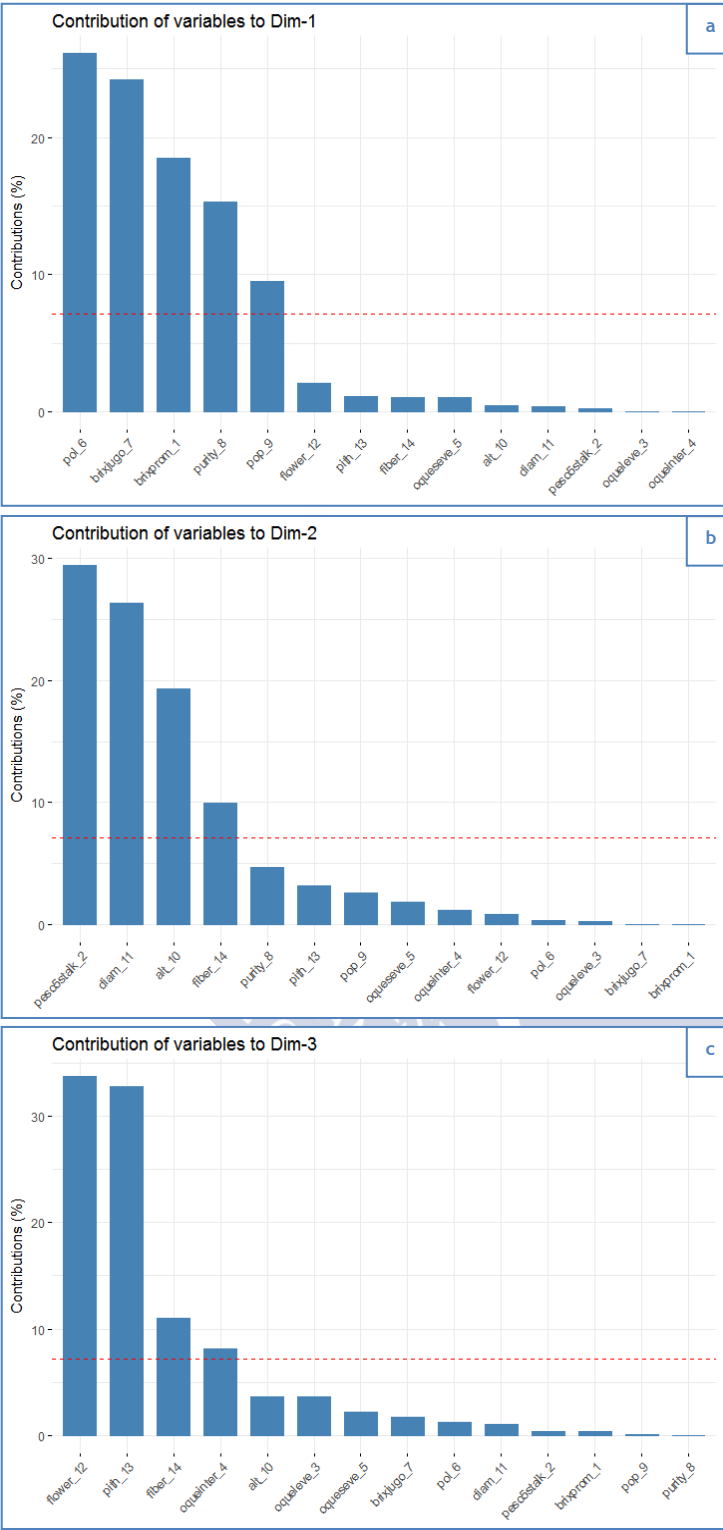


Figura 4.14. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c).

rixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oquele_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Por otro lado, los conjuntos de datos con bajos niveles de desempeño en términos de producción de biomasa fueron los experimentos de variedades **flor** evaluados en el estrato medio (Tablas 4.2 y 4.3). En estos ensayos se obtuvieron los valores más bajos de expresión fenotípica en términos de biomasa, lo que estuvo asociado a condiciones desfavorables, principalmente de suelo (ver sección de Materiales y Métodos, epígrafe 3.1). Para estos experimentos se obtuvo también la más baja tasa de selección. Los conjuntos de datos provenientes del estrato litoral no mostraron un claro patrón de agrupamiento, no obstante se apreció que tendieron a una mayor producción de biomasa y a una menor concentración de sacarosa (Tablas 4.2 y 4.3).

La variable Población por metro lineal (**pob_9**) en la CP1 tuvo dirección contraria a los vectores de las variables relacionadas con la concentración de azúcar, lo que está en concordancia con lo que se discutió anteriormente sobre una correlación atípica entre las altas poblaciones y las bajas concentraciones de azúcar (Figura 4.13).

Las tres primeras CPs acumularon el 51.21% de la varianza total (Figura 4.15), separándose en la parte positiva de la CP3 los genotipos que presentaron mayores valores relacionados con el corcho (**pith_13**), la floración (**flower_12**) y la fibra (**fiber_14**).

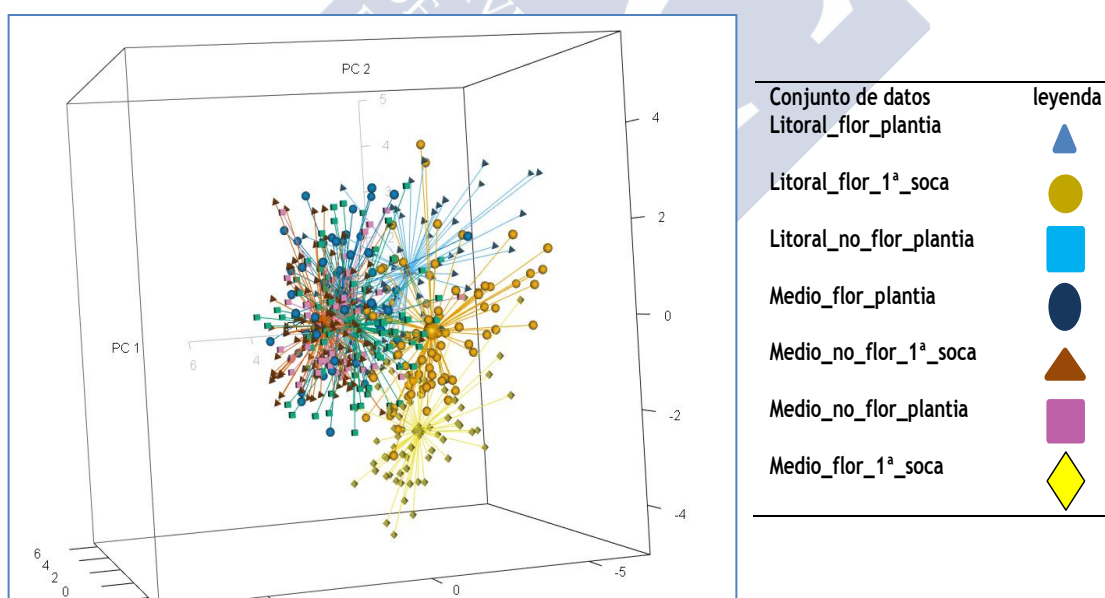


Figura 4.15. Distribución de todos los siete conjuntos de datos según tres CPs (CP1 y CP2 y CP3) para siete conjuntos de datos.

Cuando los ACP se ejecutaron para cada uno de los siete conjuntos de datos, se obtuvieron patrones similares de distribución en términos de las varianzas (*eigenvalues*) y de las cargas (*eigenvectors* o *loadings*). A partir del epígrafe 4.10.1 se presentan los resultados para tres de los conjuntos de datos más relevantes de los ensayos-

Los resultados de tres de los siete conjuntos de datos presentaron una mayor relevancia debido a que el primer conjunto de datos (**medio_flor_plantía**) fue en el que se evaluaron una menor frecuencia de genotipos con un buen desempeño fenotípico, tanto en plantía como en primera soca (Tablas 4.2 y 4.3). La mayoría de los genotipos del total en este ensayo no presentaron un buen desempeño fenotípico para las variables estudiadas, tanto en plantía como en primera soca, de ahí que tuvieran las frecuencias más bajas de genotipos evaluados (Tabla 4.2). La zona de estudio para este conjunto de datos correspondió, según se indicó en el apartado de Materiales y Métodos, a la ZAE 18.

Otro conjunto de datos (**litoral_no_flor_plantía**) se seleccionó como muestra, ya que su cultivo se realizó en condiciones diferentes. Éste se cultivó en el estrato litoral (Tabla 3.1 y Figura 3.2 de la sección de Materiales y Métodos) y, además, presentó una frecuencia intermedia de genotipos con buen desempeño fenotípico (42) (Tabla 4.2).

El otro conjunto de experimentos (**medio_no_flor_1ª_soca**) pudo ser ubicado en mejores condiciones de suelo en el estrato medio (ZAE 9) (Epígrafe 3.1 del capítulo de Materiales y Métodos). Este conjunto de datos presentó la mayor frecuencia de genotipos con buen desempeño (49, Tabla 4.2). En general, este conjunto de datos fue el que más genotipos entregó tanto en plantía como en primera soca (78 genotipos en total).

En resumen, se seleccionaron como una muestra de los siete conjuntos de datos estudiados, tres de los conjuntos de datos para la presentación de los resultados, según tres niveles:

- a) El conjunto de datos con la menor frecuencia de genotipos con buen desempeño fenotípico (30) (**medio_flor_plantía**) (Tabla 4.2).
- b) El conjunto de datos con frecuencias intermedias de genotipos con buen desempeño fenotípico (42) (**litoral_no_flor_plantía**) (Tabla 4.2).
- c) El conjunto de datos con las frecuencias más altas de genotipos con buen desempeño fenotípico (49) (**medio_no_flor_1ª_soca**) (Tabla 4.2).

La presentación de tres de los siete grupos permitirá contrastar los resultados, como una muestra consistente del resto de conjuntos de datos. Esto permitirá, a su vez, no resultar redundantes sin sufrir pérdidas importantes de información.

4.10.1 Medio Flor Plantía (**medio_flor_plantía**)

El total de la varianza para este conjunto de datos, considerando las 14 variables en estudio, fue de 14,0 (Tabla 4.10).

Tabla 4.10. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el ensayo **medio_flor_plantía**.

Componente Principal	Varianza	Proporción de la varianza	proporción acumulada
CP1	3.40	24.30	24.30
CP2	2.32	16.57	40.87
CP3	1.75	12.49	53.36
CP4	1.39	9.91	63.27
CP5	1.16	8.30	71.57
CP6	0.94	6.72	78.29
CP7	0.75	5.39	83.68
CP8	0.64	4.59	88.26
CP9	0.63	4.48	92.74
CP10	0.35	2.52	95.26
CP11	0.26	1.88	97.14
CP12	0.23	1.61	98.74
CP13	0.18	1.25	100.00
CP14	0.00	0.00	100.00
Total de la varianza	14.00	100.00	

Al igual que sucedió para los resultados globales, la CP1 presentó el 24.3% de la variabilidad total y está influida, principalmente, por las variables relacionadas con el contenido de azúcar (**brixprom_1**, **pol_6**, **brixjugo_7**, y **purity_8**), lo que indica que puede ser empleada como un índice de desempeño con respecto a estas variables (Tablas 4.10 y 4.11, Figura 4.16).

Por otro lado, la CP2 está influida por las variables relacionadas con la acumulación de biomasa, principalmente **peso5tallos_2** y **diam_11**, **fiber_14** y **oqueinter_4** (Figura 4.16 y 4.17.a). La altura y la población están más asociados negativamente a la CP1 y, en general,

presentaron una correlación negativa con el contenido de azúcar (Tabla 4.11). Las CP1 y CP2 explicaron el 40.9% del total de la variabilidad identificada (Tabla 4.10, Figura 4.17.a).

Tabla 4.11. Cargas (*Eigenvectors* o *loadings*) de cada variable original para las primeras cuatro CPs del conjunto de datos medio_flor_plantía.

Variable	CP1	CP2	CP3	CP4
brixprom_1	0.300	-0.071	0.142	-0.266
peso5stalk_2	-0.026	0.481	0.337	0.142
oqueleve_3	-0.062	0.113	0.319	-0.525
oqueinter_4	-0.065	-0.425	-0.019	0.144
oqueseve_5	0.045	-0.094	-0.088	0.684
pol_6	0.509	-0.072	0.198	0.054
brixjugo_7	0.494	-0.117	0.149	0.037
purity_8	0.432	0.041	0.265	0.075
pop_9	-0.199	0.082	0.091	0.010
alt_10	-0.274	0.005	0.340	0.031
diam_11	-0.039	0.466	0.224	0.244
flower_12	0.085	0.294	-0.465	-0.192
pith_13	0.252	0.146	-0.436	-0.134
fiber_14	-0.150	-0.457	0.210	-0.140

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Las respuestas fenotípicas de los diferentes genotipos estudiados fueron bajas en este ensayo, particularmente en el tema de acumulación de biomasa (Figura 4.2 y Tablas 4.2 y 4.3), tanto en el ciclo de plantía como en 1ª soca. Este bajo nivel de expresión finalmente se vio reflejado también en la presentación más baja de los índices de desempeño, incluidos en éstos los valores de las tres primeras CPs.

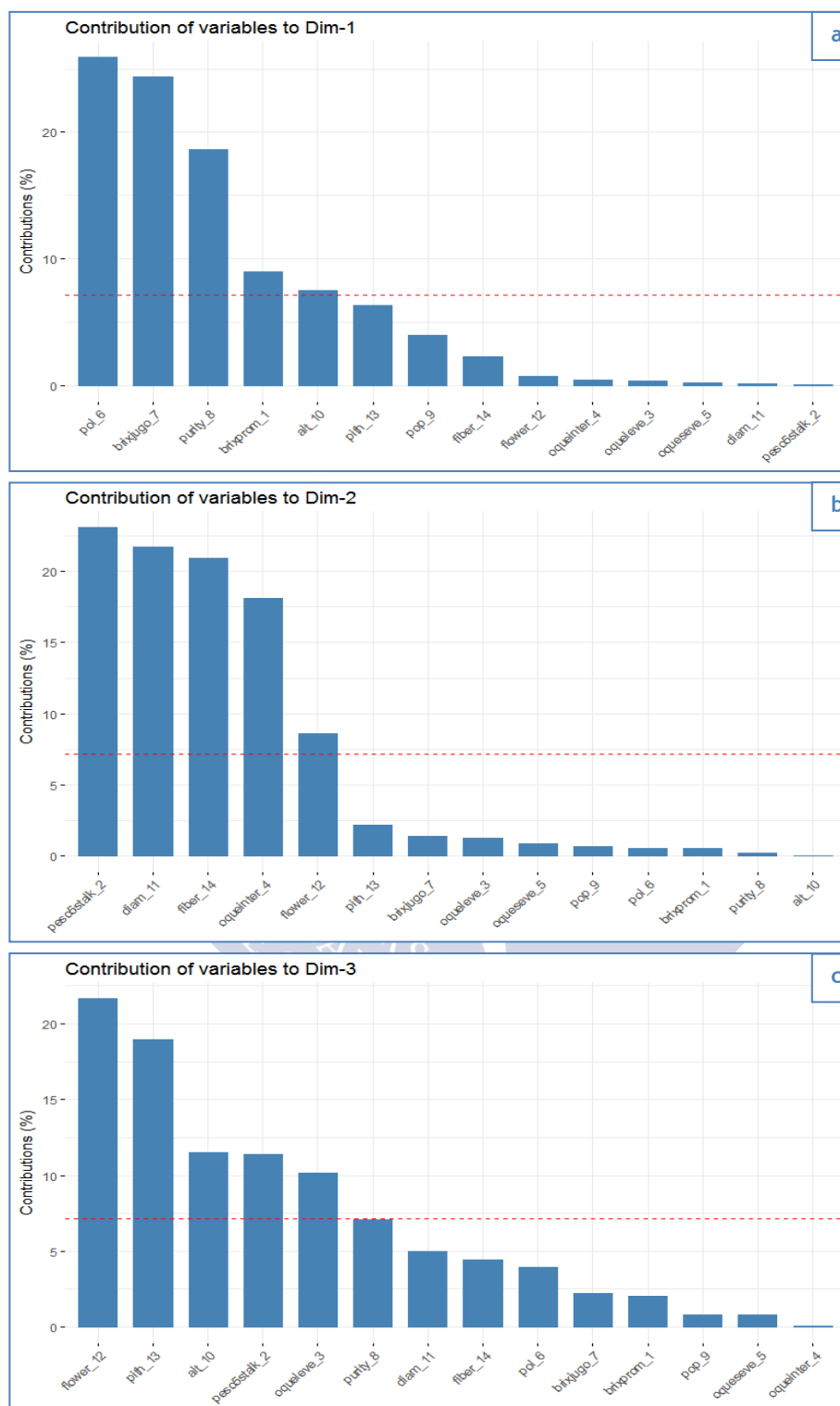


Figura 4.16. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c).

rixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueleve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Al realizar un ACP específico para el conjunto de datos **medio_flor_plantía** (Figuras 4.17.a y 4.17.b), las variables relacionadas con la concentración de azúcar tuvieron las cargas relativas más altas. En este sentido, la CP1 reflejó mayormente los efectos de las variables relacionadas con la concentración de azúcar (Figura 4.17.a), en tanto que la CP2 estuvo más asociada con las variables relacionadas con la acumulación de biomasa, en donde la variable **peso5stalk_2** fue la que presentó mayor influencia. En términos generales, las cargas más altas presentaron valores intermedios (Tabla 4.11), excepto **oqueveve_5**, la cual alcanzó una carga cercana a 0.70; sin embargo, corresponde a la CP4, la cual contribuye con poco menos del diez por ciento del total de la variabilidad (Tabla 4.10).

Al agregar la CP3 al análisis, la proporción de la variabilidad explicada fue del 53.4%. El diagrama de dispersión que representa los valores aludidos mostró que las magnitudes de los vectores (*loadings*, “cargas”, *Eigenvectors*) fueron de menor valor (Figura 4.17.b, Tabla 4.11). La CP4 estuvo asociada con Oquedad leve y severa y la proporción de la variabilidad explicada ascendió al 63.3% (Tabla 4.11). La presencia de Oquedad leve (**oqueleve_3**) estuvo relacionada con una mayor acumulación de biomasa y lo contrario ocurrió con Oquedad severa (**oqueveve_5**). En la Figura 4.16 se aprecian las contribuciones por variable en porcentaje para cada CP.

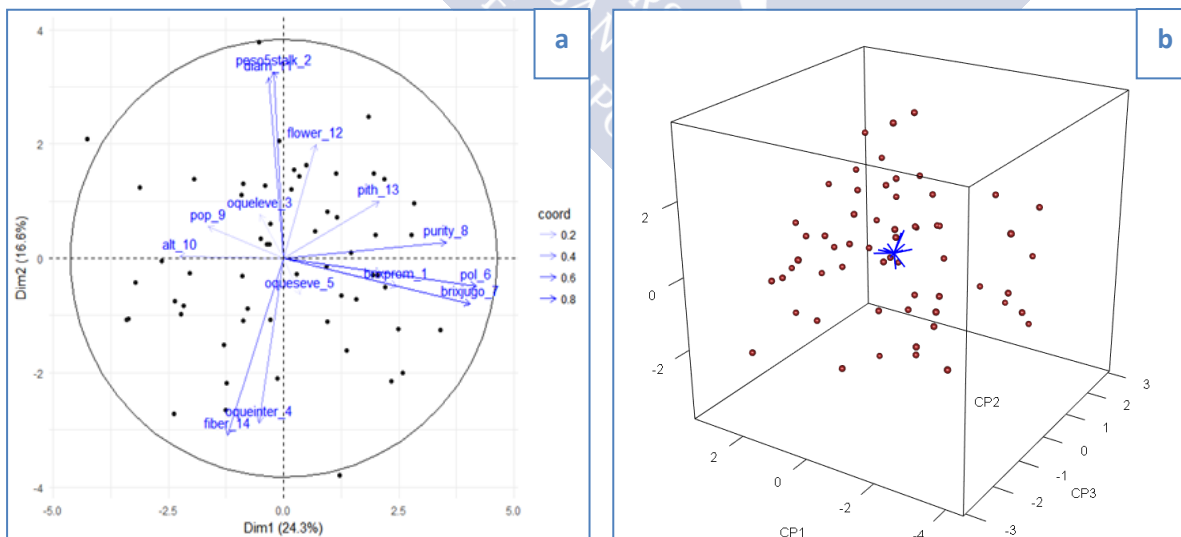


Figura 4.17. a) Gráfico De Dispersión (Biplot) CP1 y CP2; b) Gráfico De Dispersión (Triplot) CP1, CP2 y CP3 Para El Conjunto De Datos **medio_flor_plantía**.

rixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueveve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purty_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

4.10.2 Medio No Flor Primera Soca (medio_no_flor_1ª_soca)

El experimento de variedades **no_flor** del estrato medio se ubicó en la Finca “El Bálsamo”, en el lote conocido como “El Siete”. La Zona Agro-ecológica (ZAE) correspondiente fue la número 9, la cual, como se indica en el epígrafe 3.1 del capítulo de Materiales y Métodos, tiene mejores condiciones de suelo y humedad comparada con la ZAE 18, en donde se ubicó el experimento de variedades **flor** del estrato medio.

Las cuatro primeras CPs explicaron el 55.7% del total de la variabilidad encontrada en este ensayo. Para este conjunto de datos, la principal diferencia con el análisis generalizado fue que tanto la CP2 como la CP3 estuvieron relacionadas con la acumulación de biomasa, de forma negativa con la CP2 y positiva con la CP3. Ambas componentes tuvieron una participación proporcional cercana (CP2=14.9%, CP3=11.2%) en el total de la varianza, la cual alcanzó un total de 14.00 (Tabla 4.12).

Tabla 4.12. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el ensayo medio_no_flor_1ª_soca.

Componente Principal	Varianza	Porcentaje de la varianza	Porcentaje acumulado
CP1	2.78	19.85	19.85
CP2	2.09	14.92	34.76
CP3	1.57	11.20	45.96
CP4	1.37	9.76	55.72
CP5	1.14	8.16	63.88
CP6	1.00	7.17	71.05
CP7	0.80	5.69	76.75
CP8	0.73	5.19	81.94
CP9	0.65	4.67	86.60
CP10	0.56	3.99	90.59
CP11	0.51	3.67	94.26
CP12	0.47	3.36	97.62
CP13	0.31	2.22	99.83
CP14	0.02	0.17	100.00
Total de la varianza	14.00	100.00	

La CP1 acumuló el 19.8% de la variación y la CP2 el 14.9% (Tabla 4.12). Una vez más se observó que los vectores asociados con la concentración de azúcar (**pol_6**, **brixjugo_7**, **purity_8**, **brixprom_1**) estuvieron más relacionados con la CP1, y los vectores relacionados

con la biomasa se relacionaron más con la CP3 (**alt_10, diam_11, peso5stalk_2**), en tanto que las variables relacionadas con ciertos aspectos de la calidad del tallo (**pith_13, pop_9, flower_12, peso5stalk_2, oqueleve_3**) estuvieron más relacionadas con la CP2 (Tabla 4.13 y Figura 4.18 y 4.19).

Tabla 4.13. Cargas (*Eigenvectors* o *loadings*) de cada variable original con respecto a las cuatro primeras CPs para el conjunto de datos medio_no_flor_1ª_soca.

	CP1	CP2	CP3	CP4
brixprom_1	-0.403	0.016	-0.041	0.041
peso5stalk_2	0.069	-0.362	0.412	0.047
oqueleve_3	-0.046	0.322	0.134	0.266
oqueinter_4	-0.022	0.153	-0.152	0.607
oqueseve_5	-0.069	-0.067	-0.169	0.532
pol_6	-0.565	-0.042	0.034	-0.062
brixjugo_7	-0.521	0.002	0.070	-0.069
purity_8	-0.419	-0.190	-0.029	0.042
pop_9	0.058	0.442	0.282	0.184
alt_10	-0.107	-0.011	0.626	0.250
diam_11	0.152	-0.313	0.447	-0.024
flower_12	-0.019	-0.397	0.014	0.288
pith_13	0.039	-0.465	-0.256	0.110
fiber_14	-0.143	0.180	0.124	-0.2950

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

Con respecto a la dispersión con tres CPs (Figura 4.18.b), los valores de los vectores fueron relativamente similares entre todas las variables, y su influencia fue relativamente proporcional para las tres primeras CPs (Figura 4.19.b). Las contribuciones en porcentaje de las variables con respecto a las primeras tres CPs mostraron cómo los valores de las cargas, en términos porcentuales, ejercieron influencia relativa sobre cada uno de las tres primeras CPs (Figura 4. 19).

Los valores de las variables de Oquedad estuvieron más relacionadas con la CP4 (Tabla 4.13), componente de importancia relativa (9.7%) en el contexto del total de lo estudiado en este conjunto de datos.

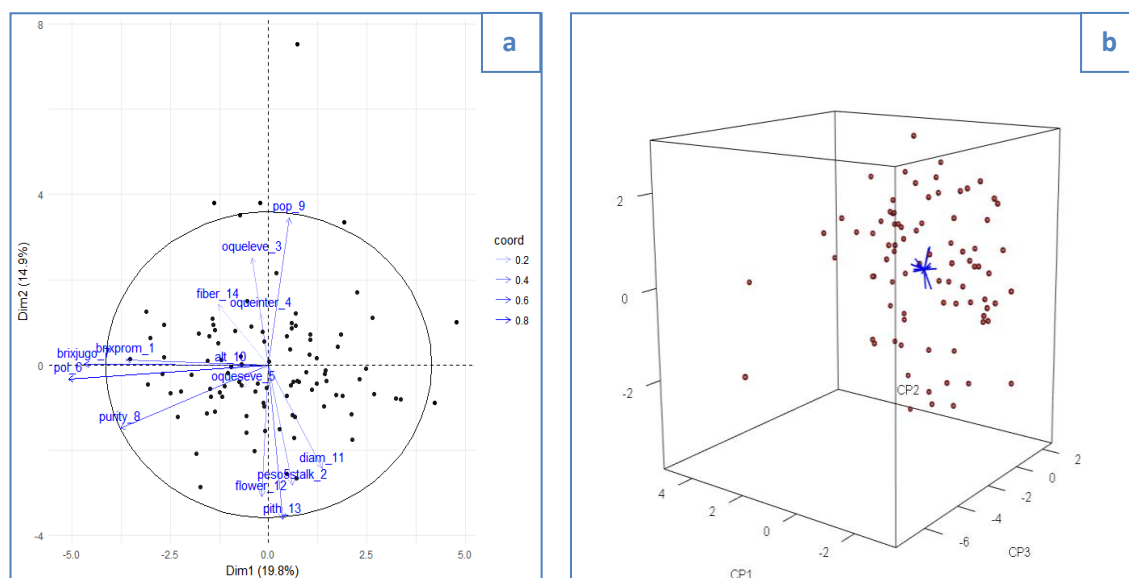


Figura 4.18. a) Gráfico de dispersión Biplot para las CP1 y CP2, y b) Gráfico de dispersión Triplot para las CP1, CP2 y CP3. Conjunto de datos medio_no_flor_1ª_soca.

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

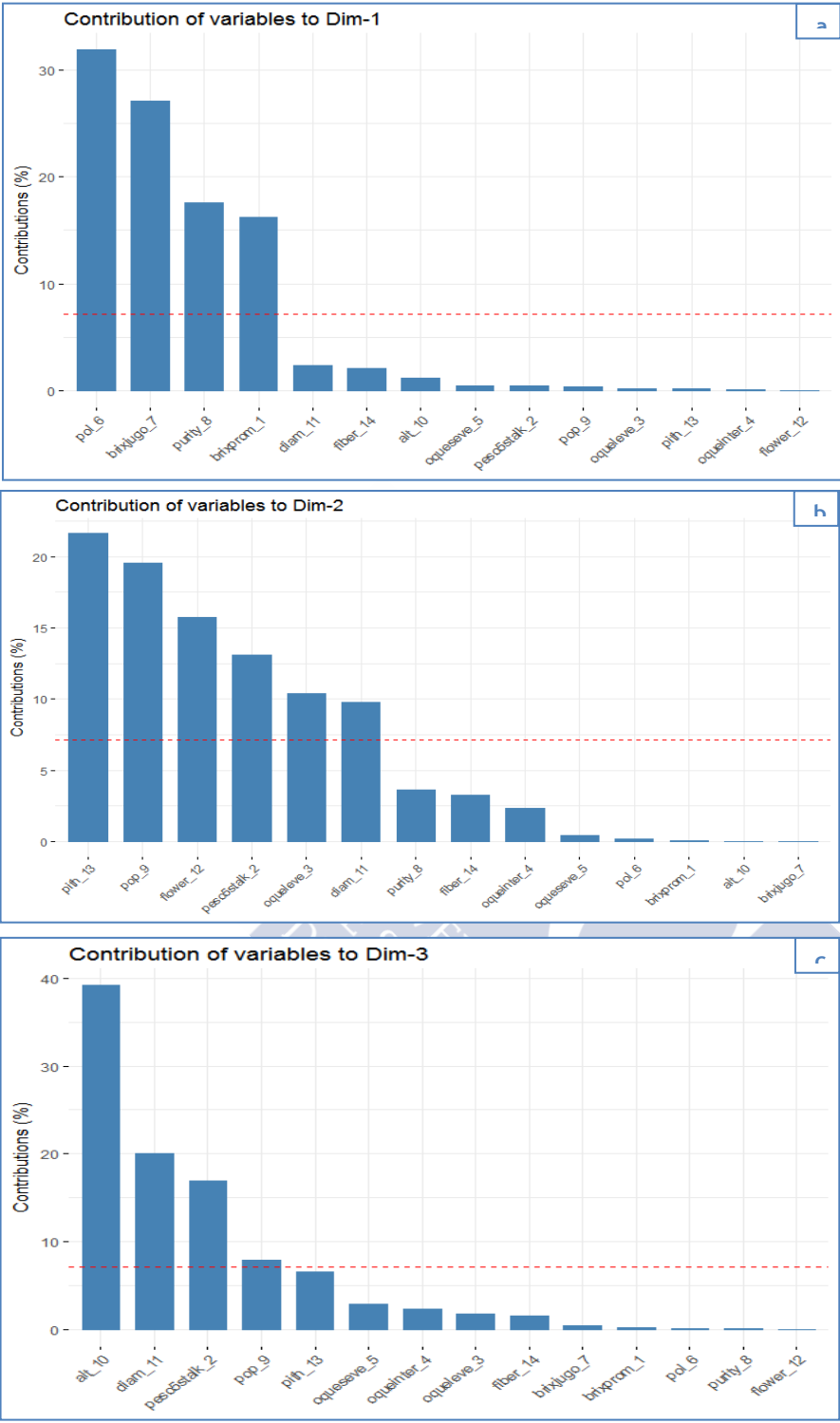


Figura 4.19. Contribución en porcentaje de las 14 variables a las CP1 (a), CP2 (b) y CP3(c) para el conjunto de datos medio_no_flor_1^a_soca.

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

4.10.3 Litoral No Flor Plantía (litoral_no_flor_plantía)

Análogamente a lo que se encontró en los otros dos conjuntos de datos analizados, la varianza total en este conjunto fue de 12.80. La variabilidad explicada por las primeras cuatro CPs fue del 58.9% (Tabla 4.14).

Tabla 4.14. Raíces características (varianzas) de cada CP, varianza total, proporción de la varianza explicada por CP y proporción acumulada para el conjunto de datos litoral_no_flor_plantía.

CP	Varianza	Proporción de la varianza	Proporción acumulada
CP1	1.74	21.63	21.63
CP2	1.43	14.64	36.27
CP3	1.37	13.34	49.62
CP4	1.14	9.32	58.93
CP5	1.03	7.58	66.52
CP6	0.96	6.59	73.11
CP7	0.95	6.48	79.59
CP8	0.87	5.38	84.97
CP9	0.84	5.02	89.99
CP10	0.70	3.45	93.44
CP11	0.64	2.91	96.35
CP12	0.57	2.31	98.65
CP13	0.41	1.17	99.82
CP14	0.16	0.18	100.00
Total de la varianza	12.80	100.00	

Una vez más, las variables asociadas a la concentración de azúcar estuvieron más relacionadas con la CP1 (Tabla 4.15, Figura 4.20); las variables asociadas a la biomasa se relacionaron, en este caso, con la CP2, si bien estas relaciones fueron de orden negativo, por lo que esta CP indica que los valores más bajos reflejan expresiones más altas con respecto a las variables en cuestión.

La CP3 estuvo más influida por las variables sobre la calidad del tallo (oquedad y corcho), así como por la población de tallos. La CP4 se asoció con algunas otras características de calidad del tallo, como la presencia de corcho (**pith_13**), y Oquedad severa (**oqueseve_5**). La variable altura (**alt_10**) también presentó una participación importante en esta componente (Tabla 4.15, Figura 4.21).

Tabla 4.15. Cargas (*Eigenvectors* o *loadings*) de cada variable original con respecto a las cuatro primeras CPs para el conjunto de datos *litoral_no_flor_plantía*.

	CP1	CP2	CP3	CP4
brixprom_1	-0.417	0.080	-0.089	0.100
peso5stalk_2	-0.149	-0.528	0.144	-0.417
oqueleve_3	-0.039	0.332	-0.055	-0.500
oqueinter_4	-0.133	0.213	0.416	-0.165
oqueseve_5	-0.130	0.082	0.480	0.029
pol_6	-0.535	0.076	-0.124	-0.071
brixjugo_7	-0.463	0.107	-0.217	-0.136
purity_8	-0.449	0.005	0.174	0.124
pop_9	-0.018	0.150	0.387	0.147
alt_10	0.062	-0.309	0.454	-0.277
diam_11	-0.153	-0.516	-0.213	-0.141
flower_12	0.135	0.144	-0.057	-0.508
pith_13	-0.136	-0.060	0.262	0.180
fiber_14	0.040	0.356	0.003	-0.300

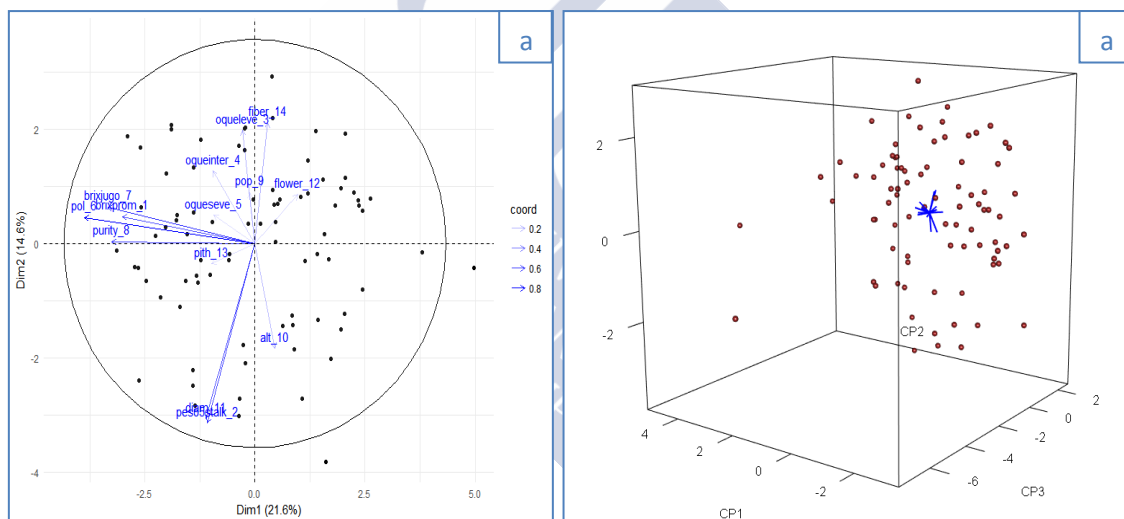


Figura 4.20. a) Gráfico de dispersión Biplot para las CP1 y CP2, y b) Gráfico de dispersión Triplot para las CP1, CP2 y CP3. Conjunto de datos *litoral_no_flor_plantía*.

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

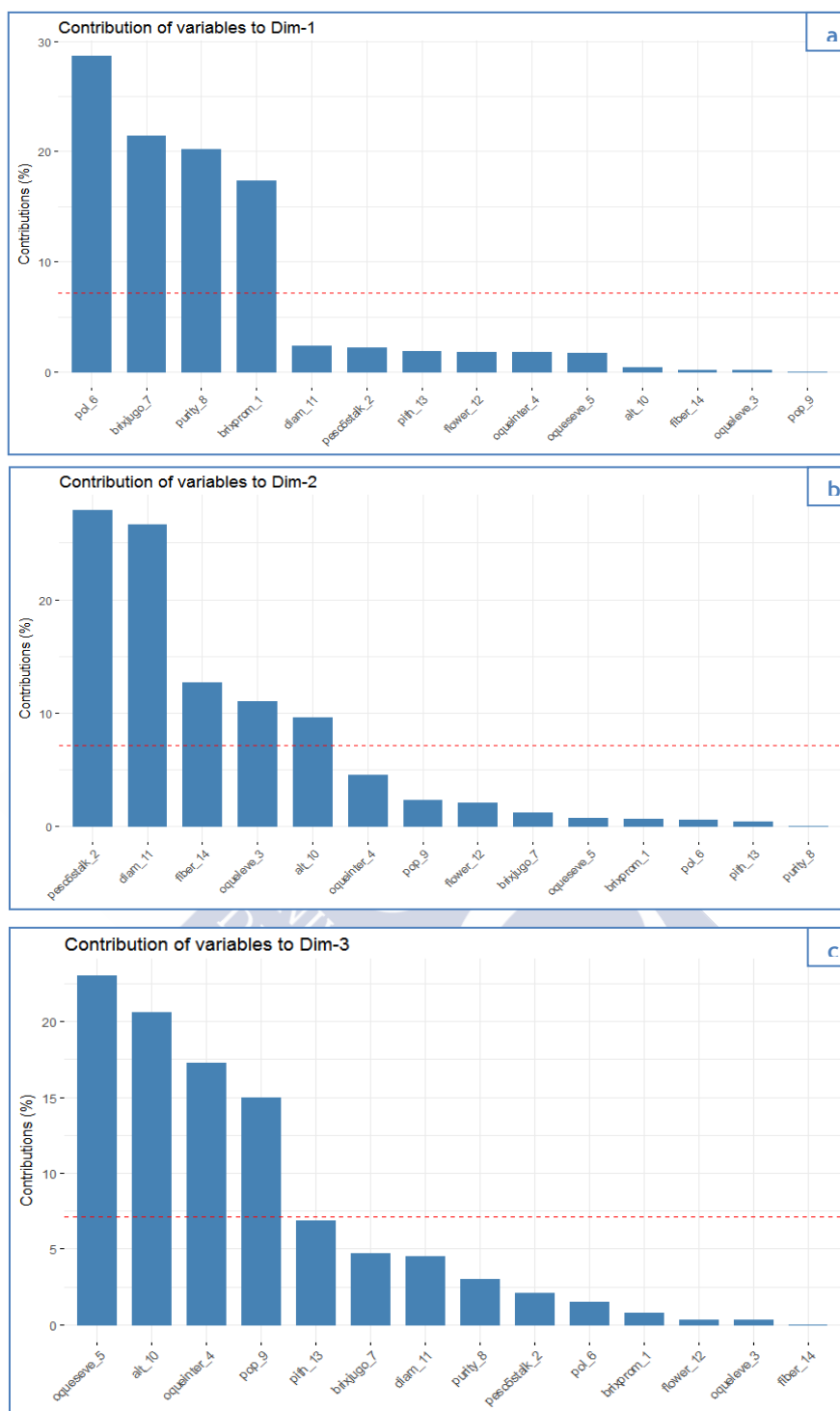


Figura 4.21. Contribución en porcentaje de las variables a las CP1 (a), CP2 (b) y CP3(c), para el conjunto de datos litoral_no_flor_plantía.

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueseve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra.

4.11 ANÁLISIS DE MODELOS LINEALES MIXTOS PARA LAS CPs PARA TODOS LOS DATOS

Se realizó un análisis de varianza combinado en el cual se consideraron todos los factores en juego según el modelo correspondiente (ver sección de Materiales y Métodos, Ecuación 3.2). Este análisis se realizó para las cuatro variables CP1, CP2, CP3 y CP4 (Tabla 4.16), considerando todas las fuentes de variación como efectos fijos. Para todas las variables en estudio se encontraron diferencias estadísticas significativas para la fuente de variación “Variedad” (ENTRY), al considerar la información de todos los siete conjuntos de datos.

Tabla 4.16. Análisis de Varianza combinado para las variables CP1, CP2, CP3 y CP4.

	CP1		CP2		CP3		CP4	
	F	Significación	F	Significación	F	Significación	F	Significación
CICLO	0	NS	0	NS	0	NS	0	NS
ESTRATO	0	NS	0	NS	0	NS	0	NS
FLORACIÓN	0	NS	0	NS	0	NS	0	NS
REPETICIÓN (REP)	0.003	NS	0.007	NS	0.147	NS	0.452	NS
VARIEDAD (ENTRY)	2.584	****	2.554	****	1.690	****	1.558	***
CICLO:ESTRATO	0.020	NS	0.058	NS	0.116	NS	0.646	NS
CICLO:FLORACIÓN	0.121	NS	0.575	NS	0.088	NS	0.958	NS
REPETICIÓN:BLOQUE INCOMPLETO	1.233	NS	0.866	NS	0.798	NS	0.402	NS
R ²	0.380		0.369		0.198		0.155	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

Tomando en consideración los resultados de la Tabla 4.16 después del análisis combinado, y para evaluar dos técnicas de análisis, se realizó también un análisis de variación para un modelo lineal mixto (Bates *et al.*, 2015) para las variables CPs. El modelo de análisis se presentó en la Ecuación 3.6 de la sección de Materiales y Métodos. Se realizó una serie de análisis de varianza para todos los conjuntos de datos para las tres primeras CPs (CP1 CP2 y CP3), siempre para los mismos tres conjuntos de datos estudiados en el epígrafe 4.10. Como ya se indicó, la CP1 estuvo altamente asociada con las características de concentración de sacarosa, en tanto que los otros dos componentes se asociaron a la acumulación de la biomasa o a calidad interna del tallo (oquedad y corcho).

Los resultados sobre normalidad de la variable CP1 para los mismos conjuntos de datos discutidos en el epígrafe 4.10, mostraron que los valores tienden a la normalidad, por lo que los resultados en los análisis posteriores son de utilidad y consistencia (Figura 4.22).

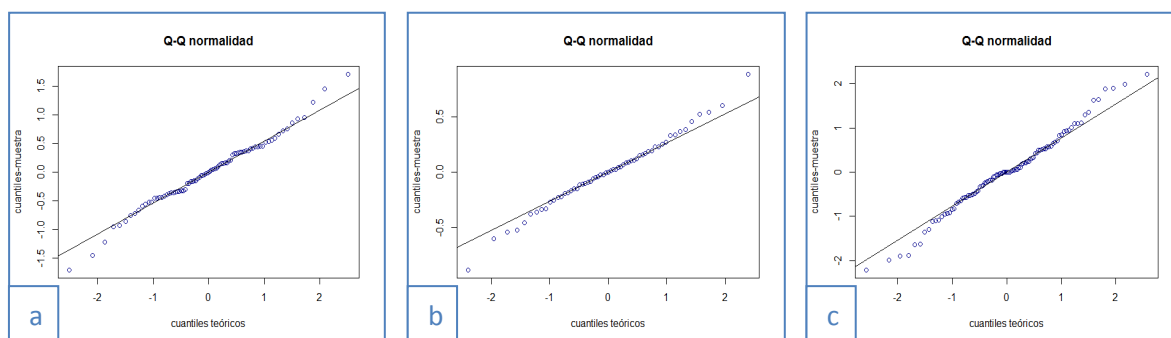


Figura 4.22. Gráfico “cuantil-cuantil” (QQ-plot) para tres conjuntos de datos: a) litoral_no_flor_plantía. b) medio_flor_plantía. c) medio_no_flor_1ª_soca para la variable (índice) CP1.

En los análisis de varianza para un modelo lineal mixto (Tabla 4.17) el factor “variedades” (ENTRY) se consideró una variable fija, y al resto de componentes del modelo como aleatorios. Todos los resultados se presentan en el Anexo 4.

Tabla 4.17. Resumen de Análisis de varianza Tipo III para 14 CPs en tres conjuntos de datos para el factor “variedad” (ENTRY). Aproximación Satterthwaite para los grados de libertad.

	medio_flor_plantía			litoral_no_flor_plantía			medio_no_flor_1a_soca		
	F	Pr(>F)	Significación	F	Pr(>F)	Significación	F	Pr(>F)	Significación
CP1	4.888	0.025	*	5.314	<0.0001	****	2.056	0.017	*
CP2	3.476	0.001	**	3.843	<0.0001	****	5.255	<0.0001	****
CP3	3.588	0.042	*	1.893	0.05	NS	3.68	<0.0001	****
CP4	0.693	0.803	NS	8.369	<0.0001	****	3.571	0.001	**
CP5	1.310	0.236	NS	4.382	<0.0001	****	3.885	<0.0001	****
CP6	1.985	0.065	NS	4.776	<0.0001	****	1.714	0.032	*
CP7	1.247	0.325	NS	3.851	<0.0001	****	1.452	0.136	NS
CP8	3.646	0.007	**	2.009	0.06	NS	1.467	0.126	NS
CP9	1.708	0.078	NS	3.714	<0.0001	****	1.648	0.086	NS
CP10	2.373	0.012	*	1.333	0.18	NS	2.696	0.003	**
CP11	1.545	0.18	NS	1.461	0.12	NS	0.909	0.629	NS
CP12	2.087	0.185	NS	1.477	0.15	NS	1.653	0.095	NS
CP13	0.474	0.975	NS	0.963	0.55	NS	1.149	0.315	NS
CP14	1.894	0.217	NS	1.275	0.26	NS	1.285	0.192	NS

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

Para los tres experimentos se detectaron diferencias significativas para la CP1 (Tabla 4.17), lo que indica que, para las variables relacionadas con la concentración de sacarosa (**brixprom_1**, **pol_6**, **brixjugo_7** y **purity_6**), se detectaron genotipos que presentaron diferencias significativas en sus promedios para la CP1 con respecto al resto de los genotipos

en evaluación. En este caso todas las “cargas” (*Eigenvectors*) entre las variables importantes relacionadas con el azúcar fueron positivas para el conjunto **medio_flor_plantía**, y negativa para los conjuntos **medio_no_flor_1ª_soca** y **litoral_no_flor_1ª_soca** (Tablas 4.11, 4.13 y 4.15). Por tanto, valores elevados de los *scores* de la CP1 identificaron valores altos de desempeño del genotipo en términos de la concentración de sacarosa para el primer conjunto, y lo contrario para los dos últimos mencionados.

Los genotipos que establecen las diferencias más significativas con respecto al grupo de variedades con el más alto desempeño en cuanto a la concentración de sacarosa (CP1) fueron CG12-337006, CG12-610018, CP72-2086 y CG12-253007 (Figura 4.23 y Tabla 4.18). Los genotipos con el más bajo desempeño, con respecto al mismo tema, (CP1) fueron CG12-146003, CG12-333017, CG12-338007, CG12-032010, y CG12-324070 (Figura 4.23 y Tabla 4.18).

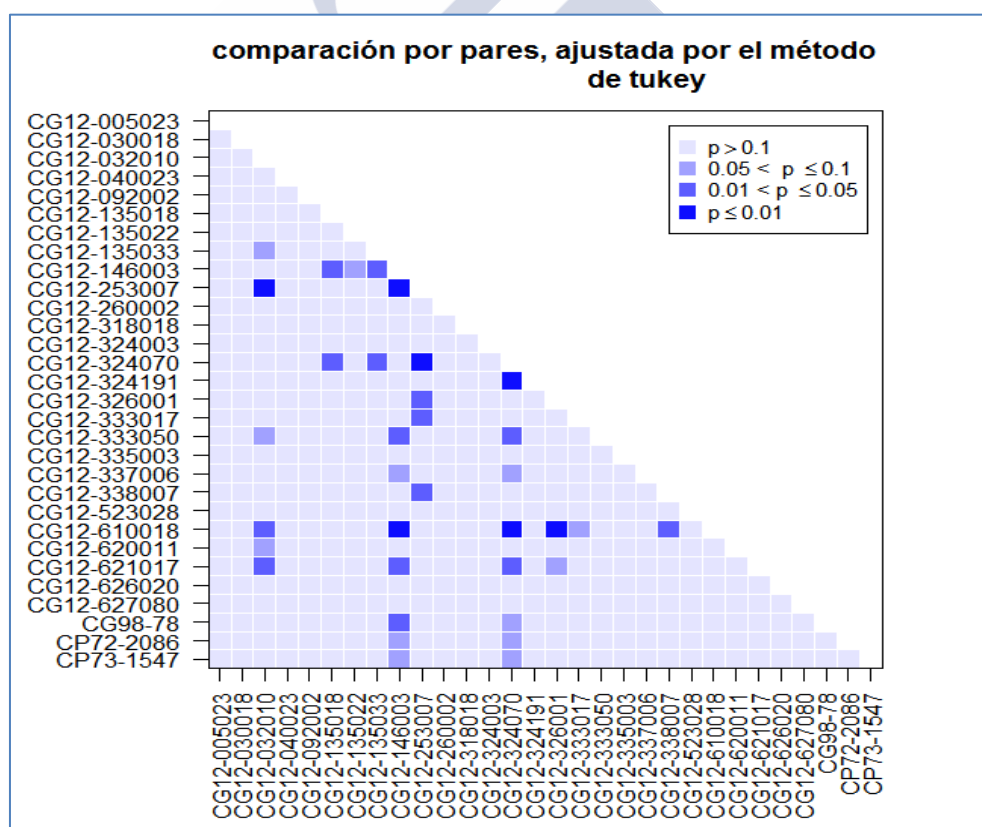


Figura 4.23. Comparación por pares ajustada por el método de Tukey para el conjunto de datos medio_flor_plantía para la CP1.

Tabla 4.18. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos medio_flor_plantía.

Variedad (ENTRY)	Agrupación por Tukey para la CP1	Media de la CP1	Media de la CP2	Media de la CP3
CG12-337006	ABCDEFGG	3.397	-1.247	0.038
CG12-610018	DG	2.465	-2.074	-1.508
CP72-2086	ABCDEFGG	2.356	0.057	-0.168
CG12-253007	D	2.344	1.721	0.988
CG12-135018	ADFG	1.597	0.212	-1.500
CG12-621017	DFG	1.271	0.463	-1.203
CP73-1547	ABCDEFGG	1.204	1.465	1.187
CG12-324191	ADEFG	1.107	0.082	1.791
CG12-620011	ABCDEFGG	0.825	1.045	1.129
CG12-135033	ADFG	0.815	-0.323	0.249
CG98-78	ACDFG	0.810	1.550	0.165
CG12-005023	ABCDEFGG	0.745	0.379	-0.893
CG12-333050	ADFG	0.735	0.668	-0.054
CG12-030018	ABCDEFGG	0.535	0.689	1.010
CG12-260002	ABCDEFGG	0.279	0.051	0.717
CG12-318018	ABCDEFGG	-0.013	-0.014	0.720
CG12-626020	ABCDEFGG	-0.216	-1.585	-1.448
CG12-523028	ABCDEFGG	-0.432	-1.296	-1.551
CG12-335003	ABCDEFGG	-0.530	3.776	-0.251
CG12-135022	ABCDEFGG	-0.536	-0.489	-1.806
CG12-326001	ABCEF	-0.584	-3.255	1.603
CG12-627080	ABCDEFGG	-0.890	0.427	-0.581
CG12-040023	ABCDEFGG	-0.896	1.206	-1.000
CG12-092002	ABCDEFGG	-1.079	-1.482	1.847
CG12-324003	ABCDEFGG	-1.085	-1.301	0.596
CG12-146003	BE	-2.077	-0.966	-1.218
CG12-333017	ABCEFG	-2.105	-0.549	-1.028
CG12-338007	ABCEF	-2.289	0.670	-0.165
CG12-032010	ABCE	-2.740	0.245	2.003
CG12-324070	BC	-3.315	-0.750	0.349

Las comparaciones entre variedades para el conjunto de datos **medio_no_flor_1ª_soca** mostraron diferencias significativas al 5% solamente entre la variedad CG12-130 con las variedades CG12-606012, CG12-512013, y una diferencia superior al 5% y menor al 10% con las variedades CG12-324187 y CG98-78, así como una diferencia superior al 10% entre las variedades CG12-606012 y CG12-606100 (Figura 4.24 y Tabla 4.19). En este caso, valores

más negativos del *score* representaron a los mejores genotipos, debido a las cargas negativas de las variables de interés (concentración de sacarosa) en la CP1.

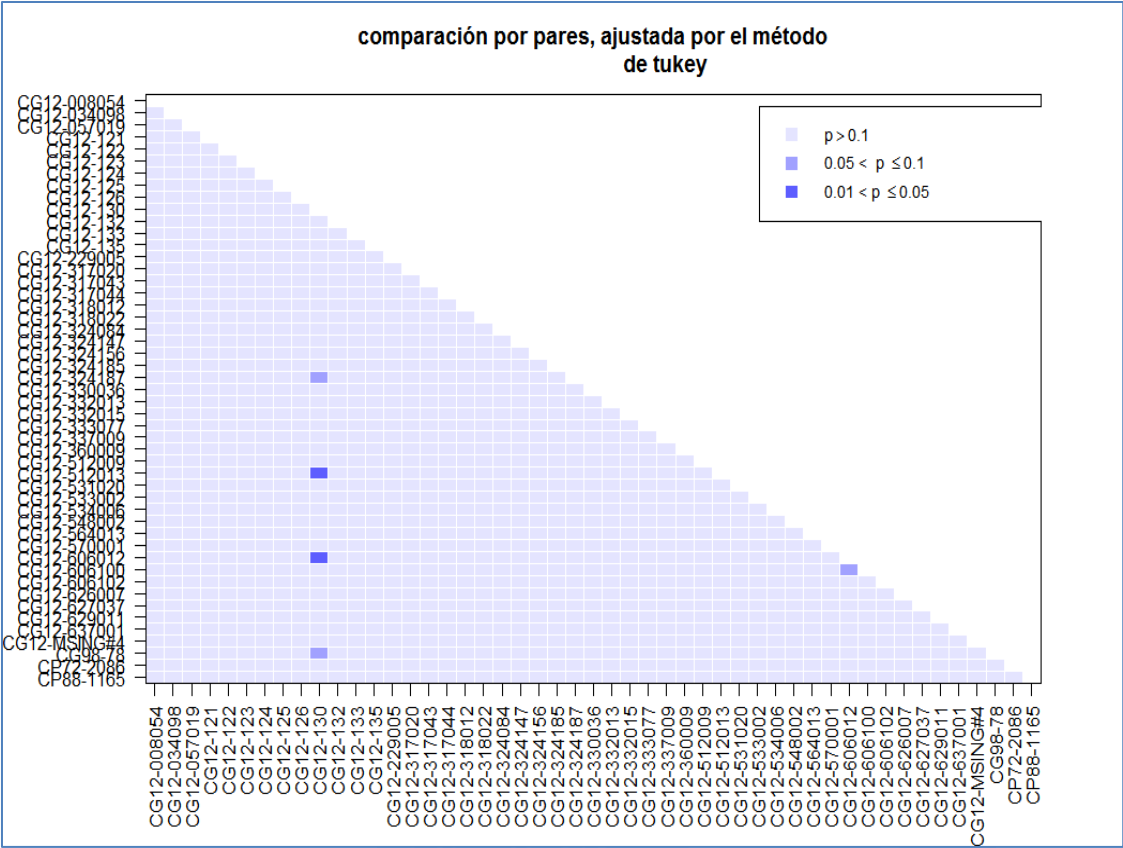


Figura 4.24. Comparación por pares ajustada por el método de Tukey para el conjunto de datos medio_no_flor_1ª_soca para la CP1.

Tabla 4.19. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos medio_no_flor_1a_soca.

Variedad (ENTRY)	Agrupación por Tukey para la CP1	Media de la CP1	Media de la CP2	Media de la CP3
CG12-130	A	3.457	-0.779	0.715
CG12-606100	AB	2.740	1.105	0.095
CG12-317044	AB	2.389	-0.892	0.970
CG12-317020	AB	1.795	-0.941	-1.979
CG12-133	AB	1.740	-0.132	0.231
CG12-564013	AB	1.703	-0.464	0.011
CG12-626007	AB	1.609	-0.234	-1.725
CG12-606102	AB	1.323	5.441	1.566
CG12-121	AB	1.241	0.949	-0.703
CG12-534006	AB	1.053	-0.645	-1.731
CG12-122	AB	0.997	-0.305	0.083
CG12-629011	AB	0.950	-0.534	0.783
CG12-229005	AB	0.879	-0.936	2.330
CG12-324084	AB	0.850	1.592	-0.277

Variedad (ENTRY)	Agrupación por Tukey para la CP1	Media de la CP1	Media de la CP2	Media de la CP3
CG12-531020	AB	0.767	0.626	-1.019
CG12-548002	AB	0.718	-0.067	-0.131
CG12-125	AB	0.590	0.919	-1.074
CP88-1165	AB	0.569	-0.553	0.374
CG12-324185	AB	0.565	0.464	-0.767
CG12-337009	AB	0.559	-2.141	0.206
CG12-318022	AB	0.332	-0.822	-0.305
CG12-324156	AB	0.260	-1.360	0.780
CG12-317043	AB	0.181	-0.851	0.502
CG12-057019	AB	0.177	-0.422	0.759
CG12-135	AB	0.108	0.754	0.818
CG12-533002	AB	0.055	0.631	0.154
CP72-2086	AB	-0.269	1.586	-1.416
CG12-637001	AB	-0.348	0.925	-0.990
CG12-570001	AB	-0.469	0.097	-1.046
CG12-126	AB	-0.562	-2.367	0.446
CG12-132	AB	-0.763	-1.385	0.763
CG12-332013	AB	-0.959	-0.304	-1.442
CG12-124	AB	-0.980	0.438	-0.922
CG12-512009	AB	-1.025	-0.784	0.592
CG12-034098	AB	-1.048	-0.561	0.389
CG12-324147	AB	-1.058	3.658	1.505
CG12-360009	AB	-1.079	0.035	-0.224
CG12-MSING#4	AB	-1.144	-2.226	2.064
CG12-123	AB	-1.226	1.175	-0.007
CG12-008054	AB	-1.270	-0.595	0.202
CG12-330036	AB	-1.282	-0.016	-0.767
CG12-333077	AB	-1.395	0.952	0.760
CG12-318012	AB	-1.400	-0.897	1.903
CG12-332015	AB	-1.407	0.264	-1.066
CG12-627037	AB	-1.413	0.304	0.861
CG12-324187	AB	-1.880	-0.224	-0.110
CG98-78	AB	-1.889	-0.895	-0.845
CG12-512013	B	-2.296	0.254	0.433
CG12-606012	B	-2.443	0.164	-1.749

Para el conjunto de datos del **litoral_no_flor_plantía** se observaron diferencias significativas para la comparación por pares entre variedades (Figura 4.25). Debido a que en este conjunto de datos también los *eigenvectors* de las variables relacionadas con el azúcar son de orden negativo en la CP1, entonces los valores más bajos de los *scores* para sus respectivos genotipos, son los que presentan las mejores características (en el Anexo 3 se presentan los *scores* para los tres conjuntos de datos discutidos y el resto de observaciones de los siete conjuntos de datos). Así es que, los mejores genotipos por su buenas propiedades en la concentración y calidad de la sacarosa fueron CG12-111, CP72-2086, CG12-113 y CG12-526031, los cuales contrastan con las variedades CG12-606129, CG12-115, CG12-339025, y CG12-110, que son las que presentaron los valores más altos de los *scores* y que, en este caso,

representaron a las variedades con un bajo desempeño en términos de concentración de sacarosa (Figura 4.25, Tabla 4.20).

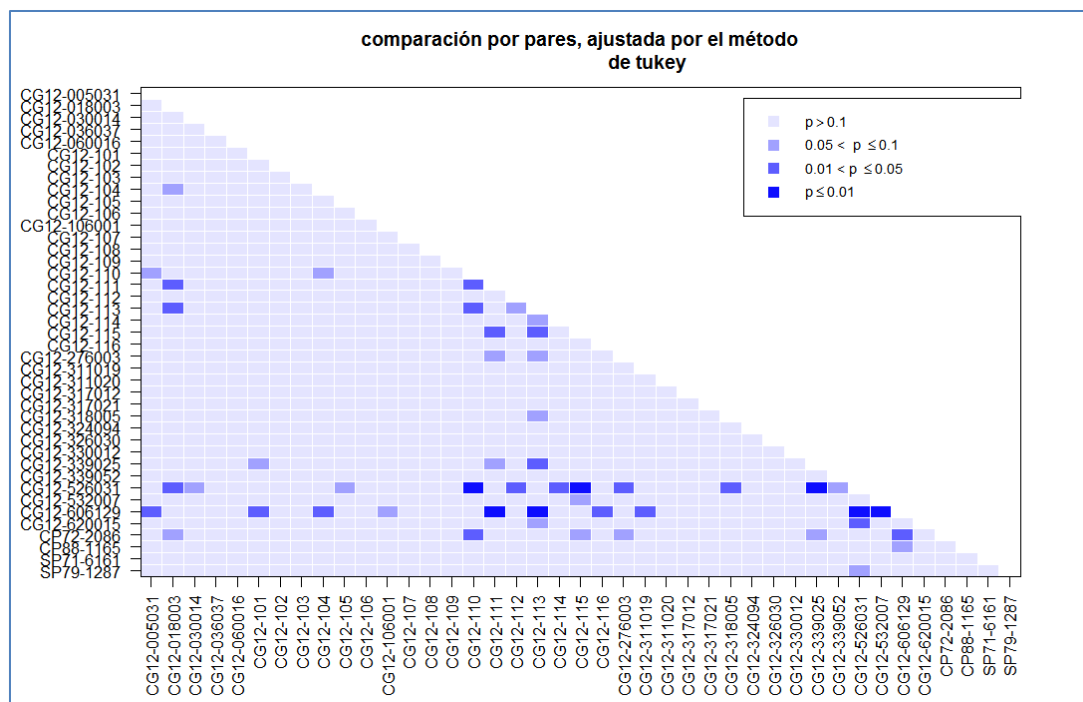


Figura 4.25. Comparación por pares ajustada por el método de Tukey para el conjunto de datos litoral_no_flor_plantía para la CP1.

Los genotipos que más diferencias presentaron fueron CG12-113 y CG12-526031, en particular con los genotipos de menor desempeño con respecto a la acumulación de sacarosa: CG12-606129, CG12-115, CG12-110, CG12-018003 y CG12-339025.

Tabla 4.20. Promedios de las tres primeras CPs y agrupamiento por el ajuste de Tukey para el conjunto de datos litoral_no_flor_plantía.

Variedad (ENTRY)	Agrupación por Tukey para la CP1	Media de la CP1	Media de la CP2	Media de la CP3
CG12-606129	F	2.926	-1.138	-1.498
CG12-115	ABF	2.234	-1.019	1.707
CG12-339025	ABDF	2.146	0.861	1.314
CG12-018003	ABF	2.097	0.390	0.514
CG12-110	AF	2.009	1.379	1.838
CG12-276003	ABCDF	1.592	0.426	-0.509
CG12-620015	ABCDF	1.575	-1.676	1.127
CG12-105	ABCDEF	1.510	-2.110	-3.269
CG12-114	ABCDF	1.421	1.384	2.060
CG12-318005	ABCDF	1.376	1.282	0.257
SP79-1287	ABCDEF	1.230	-2.542	1.065
CG12-106	ABCDEF	1.215	2.415	0.020

Variedad (ENTRY)	Agrupación por Tukey para la CP1	Media de la CP1	Media de la CP2	Media de la CP3
CG12-060016	ABCDEF	1.162	0.679	-0.337
CG12-339052	ABCDEF	1.123	0.844	0.369
CG12-030014	ABCDEF	1.098	0.505	-0.153
CG12-112	ABCDF	1.070	-0.128	0.069
CG12-103	ABCDEF	0.831	0.350	-1.312
CG12-317012	ABCDEF	0.745	-1.432	-1.987
CG12-108	ABCDEF	0.685	1.094	0.582
CG12-317021	ABCDEF	0.193	0.720	-1.753
CG12-107	ABCDEF	0.159	1.394	-0.274
CG12-102	ABCDEF	-0.304	1.671	0.041
CG12-326030	ABCDEF	-0.549	1.083	1.088
CG12-036037	ABCDEF	-0.596	0.742	-0.514
CG12-330012	ABCDEF	-0.655	0.433	-1.551
CG12-109	ABCDEF	-0.682	-1.631	-0.427
SP71-6161	ABCDEF	-0.784	0.044	1.117
CG12-311020	ABCDEF	-0.817	-2.301	-0.338
CG12-324094	ABCDEF	-0.893	-2.616	-0.590
CP88-1165	ABCDEF	-1.359	-1.586	0.587
CG12-311019	ABCDE	-1.430	-1.261	0.731
CG12-106001	ABCDEF	-1.455	-0.613	0.434
CG12-005031	ABCDE	-1.646	1.304	0.045
CG12-116	ABCDE	-1.685	-0.620	0.310
CG12-104	ABCDE	-1.723	1.079	-0.131
CG12-101	ABCDE	-1.830	0.732	-0.292
CG12-532007	ABCDE	-2.017	0.760	0.118
CG12-111	CDE	-2.139	-0.236	0.611
CP72-2086	BCDE	-2.216	-0.996	0.343
CG12-113	CE	-2.686	0.610	-0.823
CG12-526031	E	-2.934	-0.272	-0.587

4.12 ANÁLISIS MULTIDIMENSIONAL DE VARIANZA (MANOVA)

Si bien los análisis por variable son informativos, y con los índices es posible realizar una ordenación de genotipos según las funciones de cada índice, también es importante conocer las respuestas conjuntas tomando en consideración todas las variables con respecto al modelo general propuesto de análisis para cada conjunto de datos. El propósito principal de hacer un análisis de variación para un esquema multidimensional es establecer si la variabilidad representada en el vector de medias del conjunto de las variables dependientes es una función de las variables independientes planteadas (ver en la sección de Materiales y Métodos). Finalmente, la intención en estos análisis es evaluar las respuestas de los genotipos en términos de todas las variables en estudio.

Por esta razón se presentaron los principales resultados de análisis multidimensionales de varianza (MANOVA), donde la variable de respuesta fue un vector de medias, el cual incluyó

las catorce medias de las variables de respuesta respectivas originalmente propuestas (los detalles metodológicos fueron presentados en la sección de Materiales y Métodos). Para el caso de los conjuntos de datos de materiales **no_flor**, se consideraron solamente 13 variables, puesto que la variable Floración no funciona como tal, ya que, en general, los valores de “porcentaje de floración” son virtualmente cero.

De los resultados obtenidos, se presentaron las relaciones más relevantes para el total de los datos en los siete conjuntos de datos (Figuras 4.26, 4.27 y 4.28). Las variables relacionadas con la concentración de sacarosa (Figura 4.26), para todos los conjuntos de datos, mostraron una tendencia de correlación positiva entre ellas. La variable de °Brix Campo (**brixprom_1**) fue la que presentó menores valores de correlación con el resto de variables relacionadas con la sacarosa (Figura 4.26); no obstante, los valores de correlación pueden considerarse relativamente altos (Figura 4.26 y Tabla 4.5). Por otro lado, los gráficos de densidades (en la diagonal del centro de la gráfica) muestran una tendencia a la normalidad de las cuatro variables en cuestión. Los niveles de significación estadística de estas regresiones se presentaron en la Tabla 4.5.

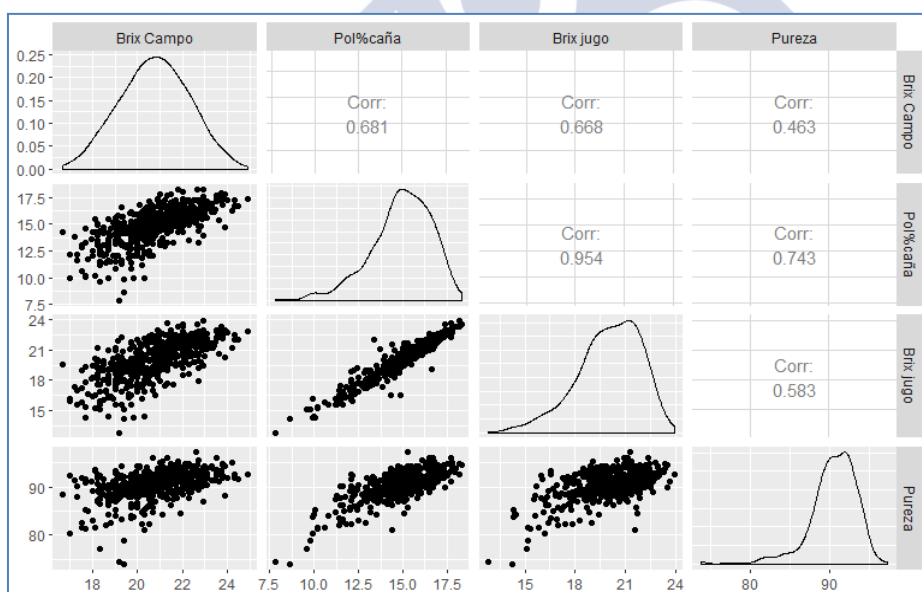


Figura 4.26. Matriz de diagramas de dispersión y “r” de Pearson entre cuatro variables relacionadas con la concentración de sacarosa para el total de datos (Los niveles de significación estadística de estas regresiones se presentaron en la Tabla 4.5.).

Para el caso de las variables relacionadas con la acumulación de biomasa (Figura 4.27) se observó que las distribuciones eran de tendencia normal, con la excepción de la variable fibra que tuvo una tendencia más bien bimodal. Para las variables en cuestión, se apreció que las

correlaciones fueron bajas, excepto para las variables Diámetro (**diam_11**) y Altura (**alt_10**), que guardan correlaciones más elevadas con la variable Peso de 5 Tallos (**peso5stalk_2**) ($r=0.58$ y 0.49 , respectivamente).

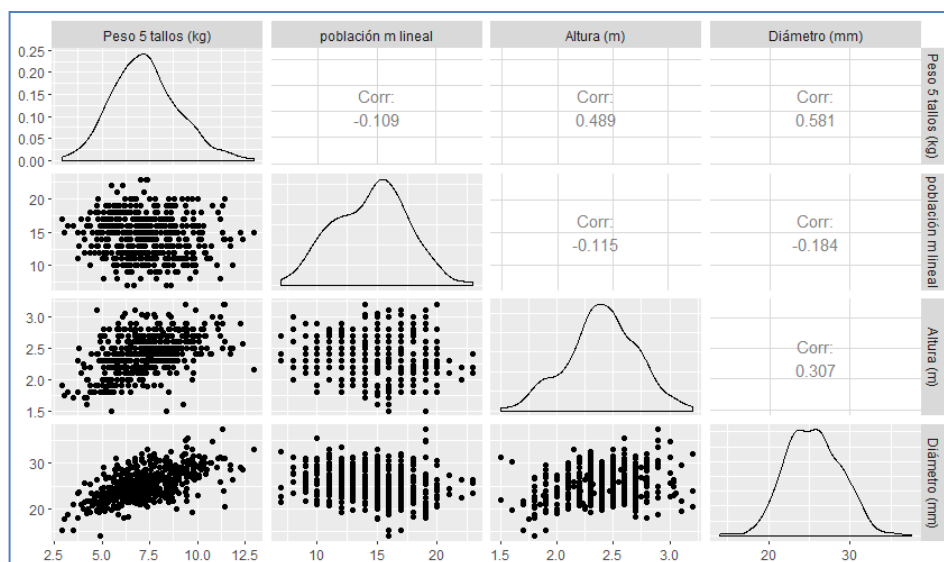


Figura 4.27. Matriz de diagramas de dispersión y “r” de Pearson entre cuatro variables relacionadas con acumulación de biomasa para el total de datos (los niveles de significación estadística se presentaron en la Tabla 4.5).

Algo similar se ha encontrado para las variables relacionadas con la calidad interna del tallo (Figura 4.28). En este caso es importante recordar las transformaciones de las variables oquedad (**oqueleve_3**, **oqueinter_4** y **oqueleve_5**), corcho (**pith_13**) y floración (**flower_12**) (ver en la sección de tratamiento de las variables en el apartado de Materiales y Métodos, epígrafe 3.5). Por ejemplo, la correlación entre floración (**flower_12**) y fibra (**fiber_14**), aparece como negativa (-0.20), cuando la correlación en realidad es positiva.

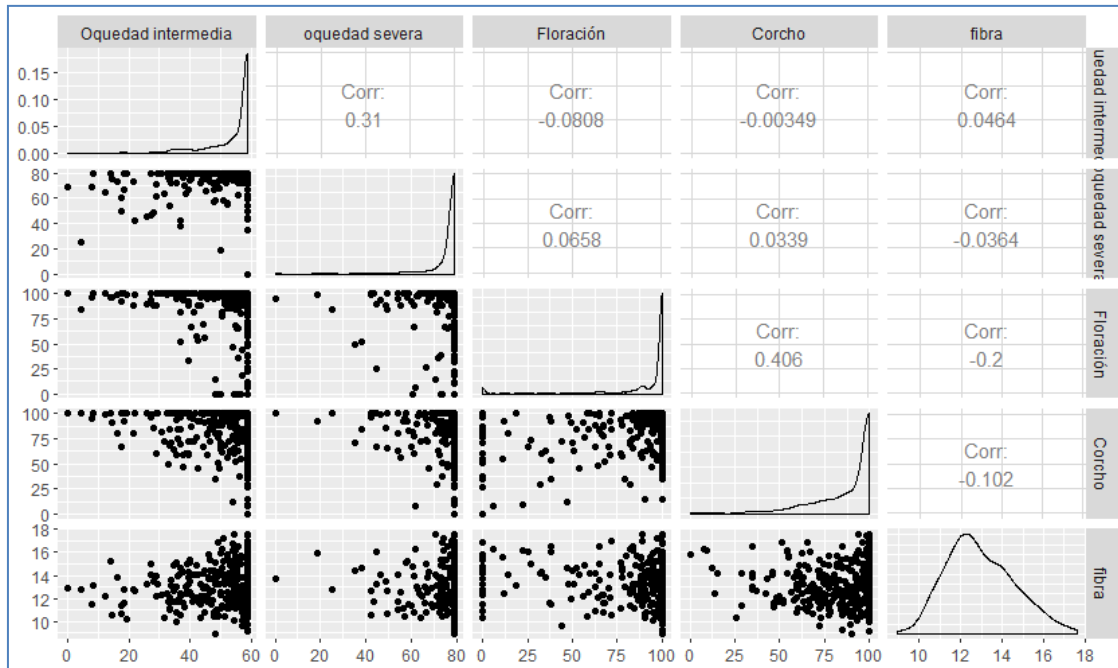


Figura 4.28. Matriz de diagramas de dispersión y "r" de Pearson entre cinco variables relacionadas con características de calidad del tallo para el total de datos (significación estadística en la Tabla 4.5).

La respuesta multidimensional analizada mediante MANOVA para cada uno de los tres conjuntos de datos que se han venido estudiando a lo largo de este documento (**medio_flor_plantía**, **medio_no_flor_1ª_soca** y **litoral_no_flor_plantía**) (Tabla 4.21) mostraron diferencias significativas para el estadístico F (Carey, 1998; Hair *et al.*, 20104), siendo analizados los cuatro diferentes valores del estadístico F (calculado a partir del estadístico A) (Carey, 1998) que ofrece **R**: Pillai, Wilks, Hotelling-Lawley y Roy. El estadístico "F" de Pillai ha demostrado ser el estadístico más robusto para el modelos como el propuesto (Carey, 1998). En todo caso, las cuatro aproximaciones mostraron diferencias estadísticas

Tabla 4.21. Resumen de los Análisis Multidimensionales de Varianza (MANOVA). Solo para la variable de clase (independiente) VARIEDAD (ENTRY).

Conjunto de datos	Método	Valor en el Método	F aproximada	Pr (>F)	Significación
medio_flor_plantía	Pillai	8.78	1	0.005	**
	Wilks	< 0.0001	2	0.0004	***
	Hotelling-Lawley	1745.00	-4	NA	NA
	Roy	1326.00	503	1.21E-13	****
medio_no_flor_1ªsoca	Pillai	8.67	1	0.0004	***
	Wilks	< 0.0001	2	9.56E-06	****
	Hotelling-Lawley	70.00	2	4.80E-06	****
	Roy	21.00	18	4.28E-11	****
litoral_no_flor_plantía	Pillai	9.32	2	6.05E-14	****
	Wilks	< 0.0001	3	2.00E-16	****
	Hotelling-Lawley	93.90	3	2.00E-16	****
	Roy	22.50	18	2.19 ⁻¹³	****
litoral_flor_1ªsoca	Pillai	10.07	2	2.00E-16	****
	Wilks	<0.00001	4	2.00E-16	****
	Hotelling-Lawley	175.80	6	2.00E-16	****
	Roy	64.90	58	2.00E-16	****
litoral_flor_plantía	Pillai	8.67	1	0.0004	***
	Wilks	<0.00001	2	9.56E-06	****
	Hotelling-Lawley	70.00	2	4.80E-06	****
	Roy	21.00	18	4.28E-11	****
medio_flor_1ªsoca	Pillai	9.32	2	6.05E-14	****
	Wilks	<0.00001	3	2.00E-16	****
	Hotelling-Lawley	93.90	3	2.00E-16	****
	Roy	22.50	18	2.19E-13	****
medio_no_flor_plantía	Pillai	10.29	2	4.65E-07	****
	Wilks	<0.00001	2	2.18E-08	****
	Hotelling-Lawley	438.00	4	1.38E-06	****
	Roy	280.00	180	2.00E-16	****

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

Se encontraron, para cada conjunto de datos, diferencias altamente significativas para los cuatro estadísticos obtenidos en la fuente de variación genotipos (Variedades, ENTRY). No se realizaron pruebas *post hoc*, ya que para un MANOVA no se cumplen condiciones para la realización de pruebas discrecionales de separación de medias. Por otro lado, las comparaciones de genotipos por contrastes planificadas *a priori*, son complejas y muy complicadas de realizar. En todo caso, los MANOVAs son una base importante para establecer que, para el vector de medias de las 14 variables en estudio, las principales diferencias encontradas entre genotipos son debidas a un efecto genético.

4.13 ÍNDICES DE SELECCIÓN (DESEMPEÑO)

Se obtuvieron resultados para siete conjuntos de datos, según se apuntó en el apartado de Materiales y Métodos. Se presentaron los resultados de tres de estos conjuntos, los cuales corresponden con los tres conjuntos de datos que se han venido discutiendo hasta aquí, según la justificación que se presentó en el epígrafe 4.10. Los valores correspondientes a todos los índices, en todos los conjuntos de datos, se presentan en el Anexo 3.

4.13.1 Índices de selección para el conjunto de datos medio flor plantía (medio_flor_plantía)

4.13.1.1 Análisis de correlación entre índices

Si bien muchos de los cálculos entre índices son de naturaleza diferente, algunos de ellos presentaron similitudes según los antecedentes teóricos de los cuales se derivan. Así, los índices derivados de un ACP (*Eigenanalysis*), como ESIM y RESIM, guardaron relación con las CPs derivadas directamente de los conjuntos de datos originales. Los índices que contienen restricciones para su cálculo (KN, RESIM) también guardaron ciertas relaciones, debido a que la restricción para el cálculo del índice fue con base en las mismas variables para ambos índices.

Los valores de correlación entre los índices mostraron que en general hay cierta relación entre varios grupos de ellos, así los basados en el ACP (*Eigenanalysis*) están más relacionados entre sí y BLUP y TAH, como es lógico debido a que uno se deriva del otro, muestran altos niveles de correlación (Tabla 4.22). En otros casos como Smith y KN, presentaron variaciones en términos de sus relaciones con el resto y entre sí. En general, una mayor expresión fenotípica derivó en correlaciones más claras entre los índices como se verá adelante.

Tabla 4.22. Valores de correlación r de Pearson y probabilidad de F para probar la hipótesis $\beta_1 = 0$ para nueve índices de desempeño en el conjunto de datos medio_flor_plantia.

	TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3
TAH		****	NS	NS	NS	NS	NS	****	**
BLUP	0.99		NS	NS	NS	NS	NS	****	**
ESIM	-0.04	-0.03		****	*	NS	NS	NS	NS
RESIM	-0.05	-0.03	1		**	NS	NS	NS	NS
KN	0.15	0.15	-0.41	-0.46		****	NS	NS	NS
Smith	0.22	0.21	-0.06	-0.11	0.82		NS	NS	NS
CP1	0.03	0.07	0.20	0.20	0.25	0.22		NS	NS
CP2	0.69	0.68	0.09	0.09	0.02	0.13	0.11		NS
CP3	0.50	0.52	0.08	0.09	-0.21	-0.23	-0.05	0.08	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

4.13.1.2 Índice “Smith” y “Restrictive Kempthorne and Nordskog Selection Index” (Índice KN)

En la Tabla 4.23 se presentan los valores promedio para las variables en estudio por genotipo en evaluación y el índice producto de las funciones de Smith (1936), Kempthorne y Nordskog (1959) y el TAH, ordenados descendientemente según los valores del índice de Smith.

Tabla 4.23. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_flor_plantía, ordenados descendientemente según los valores del índice de Smith.

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14	Smith	KN	TAH
CG98-78	8.4	77.2	44.1	77.1	21.7	15.3	20.7	89.7	10.8	2.6	27.5	91.3	48.0	12.9	1.056	1.119	25.1
CG12-253007	8.2	82.3	58.8	77.4	20.1	15.4	20.8	90.2	12.5	2.7	26.2	61.6	41.9	13.0	0.896	1.367	24.7
CG12-146003	7.3	71.1	47.8	78.0	22.3	16.5	22.0	91.0	10.0	2.6	24.1	70.6	63.2	13.6	0.887	0.356	13.8
CP72-2086	9.0	71.5	43.4	77.1	22.0	16.2	21.5	92.0	9.5	2.6	29.9	56.5	63.0	12.3	0.806	0.407	23.5
CG12-620011	8.7	79.8	50.2	68.0	21.3	15.8	21.1	91.3	12.3	2.5	27.1	88.9	83.5	12.3	0.765	0.189	24.7
CG12-627080	7.5	80.5	56.9	75.0	21.5	15.3	20.6	90.5	12.3	2.5	25.5	88.8	59.7	12.8	0.603	0.539	17.9
CG12-621017	8.6	72.2	52.8	66.8	21.6	15.6	20.9	90.6	11.3	2.6	27.8	70.2	85.0	13.8	0.564	0.670	15.5
CG12-135018	9.4	75.4	51.4	81.9	20.7	15.2	20.4	90.3	12.3	2.5	28.3	55.8	64.8	13.2	0.412	0.685	19.1
CG12-092002	9.3	84.1	51.4	76.7	20.7	15.2	20.6	89.9	12.3	2.5	28.1	69.7	61.7	13.2	0.364	0.563	21.7
CG12-135022	9.7	72.3	29.0	66.7	20.5	14.9	20.1	90.2	11.5	2.4	28.3	95.0	79.5	11.3	0.318	0.201	14.8
CG12-333017	7.5	73.1	47.3	78.6	21.4	15.1	20.7	88.7	12.5	2.5	26.3	95.5	48.1	12.3	0.274	0.075	17.2
CG12-326001	8.0	79.9	43.4	65.6	21.7	16.1	21.7	90.1	11.3	2.4	27.8	90.3	72.4	12.9	0.260	-0.238	16.5
CG12-135033	9.8	63.1	29.0	71.8	20.5	14.6	19.7	89.9	11.5	2.4	29.6	98.5	92.0	11.3	0.212	0.364	14.6
CG12-030018	6.6	72.9	55.9	80.4	21.0	15.5	21.0	90.1	12.0	2.6	25.4	50.9	72.9	13.8	0.210	0.773	25.3
CG12-335003	7.5	68.7	25.4	71.6	21.6	16.8	22.1	92.3	11.8	2.3	27.3	90.5	80.9	11.6	0.027	-1.051	25.9
CG12-523028	7.5	77.2	44.4	65.1	20.5	14.9	20.0	90.1	13.3	2.5	26.5	86.8	74.4	13.1	-0.101	0.492	14.2
CG12-610018	7.6	80.8	47.3	63.7	20.2	16.2	21.6	91.2	12.5	2.4	27.0	73.9	78.3	13.0	-0.169	-0.054	11.6
CG12-040023	8.1	76.2	46.6	77.3	21.7	15.0	20.6	88.5	13.0	2.3	27.4	86.3	67.9	13.9	-0.241	-0.198	18.5
CG12-338007	7.4	52.6	50.2	58.2	21.5	15.7	20.9	91.4	11.5	2.4	25.0	89.0	77.3	14.6	-0.242	0.173	19.2
CG12-324003	5.9	57.1	47.5	79.3	20.5	14.7	20.0	89.3	12.5	2.5	23.7	88.0	78.2	12.8	-0.266	0.548	15.5
CG12-626020	7.7	73.6	55.9	66.3	21.7	16.0	21.1	91.9	11.8	2.3	25.4	82.3	87.8	14.0	-0.270	-0.412	12.0
CP73-1547	7.6	81.8	57.6	79.3	21.2	16.2	21.4	92.3	10.0	2.3	29.5	69.5	84.0	11.8	-0.354	-0.477	21.1
CG12-005023	6.5	81.4	55.9	75.3	21.0	16.3	21.6	91.7	12.0	2.4	24.5	58.1	79.6	13.8	-0.365	-0.242	16.1
CG12-333050	6.4	70.1	42.0	76.7	20.8	15.5	21.1	89.8	11.3	2.3	24.4	82.8	68.2	11.9	-0.416	-0.366	21.9
CG12-337006	7.7	63.8	40.1	59.1	21.7	16.2	21.5	91.4	10.8	2.3	26.6	98.8	84.2	13.9	-0.441	-0.521	11.4
CG12-318018	6.7	48.0	58.8	78.8	20.8	16.1	21.5	91.4	12.0	2.2	26.5	90.9	91.3	12.5	-0.810	-0.877	20.2
CG12-032010	8.1	75.6	46.6	72.2	21.7	15.1	20.6	89.0	13.0	2.2	27.6	74.2	66.6	13.9	-0.818	-0.809	23.9
CG12-324191	6.3	78.7	53.8	64.7	21.7	16.6	22.3	90.8	9.8	2.2	25.0	74.7	88.2	13.6	-0.996	-1.210	15.4
CG12-260002	7.8	58.9	58.8	76.9	20.0	15.1	20.4	90.1	12.0	2.2	28.0	85.3	73.7	11.6	-1.034	-0.673	23.1
CG12-324070	5.3	66.4	46.3	77.6	21.5	16.0	21.9	89.1	11.8	2.1	23.0	96.1	93.6	13.2	-1.132	-1.398	15.0

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

Al realizar las comparaciones entre los genotipos en evaluación, las propuestas para selecciones fueron diferentes en ambos índices (Tabla 4.24). Los genotipos priorizados por Smith fueron diferenciales en al menos un 50% con respecto a los priorizados por KN.

Tabla 4.24. Resumen de las Salidas de RINDSEL para los índices Smith y KN y la jerarquización de la fracción superior (10%).

Índice	Estadístico	Valor del estadístico	Fracción superior (10%)	Valor del índice
Smith	Covarianza entre el índice y el VMG	17.718	CG98-78	1.056
	Varianza del índice	0.110	CG12-253007	0.896
	Varianza del VMG	451.724	CG12-146003	0.887
	Correlación entre el índice y el VMG	0.999	CP72-2086	0.806
KN	Covarianza entre el índice y el VMG	2.214	CG12-253007	1.367
	Varianza del índice	0.038	CG98-78	1.119
	Varianza del VMG	451.724	CG12-030018	0.773
	Correlación entre el índice y el VMG	0.534	CG12-135018	0.685

VMG: Valor de la Mejora Genética (*Breeding value*).

4.13.1.3 *Eigenanalysis-based Selection Index Method* (ESIM) y *Restrictive Eigenanalysis-based Selection Index Method* (RESIM)

En la Tabla 4.25 se presentan los resultados para los índices **ESIM** y **RESIM** conjuntamente, ya que en todos los casos la correlación entre ambos fue muy alta, y sus resultados en términos de selección fueron prácticamente los mismos ($r = 1$). El valor del TAH se consignó como un valor de referencia agronómica sobre el desempeño de los genotipos.

Tabla 4.25. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_flor_plantía.

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14	ESIM	RESIM	TAH
CG98-78	8.4	77.2	44.1	77.1	21.7	15.3	20.7	89.7	10.8	2.6	27.5	91.3	48.0	12.9	2.30	2.41	25.1
CG12-253007	8.2	82.3	58.8	77.4	20.1	15.4	20.8	90.2	12.5	2.7	26.2	61.6	41.9	13.0	1.18	1.31	24.7
CG12-146003	7.3	71.1	47.8	78.0	22.3	16.5	22.0	91.0	10.0	2.6	24.1	70.6	63.2	13.6	-1.64	-1.44	13.8
CP72-2086	9.0	71.5	43.4	77.1	22.0	16.2	21.5	92.0	9.5	2.6	29.9	56.5	63.0	12.3	1.40	1.43	23.5
CG12-620011	8.7	79.8	50.2	68.0	21.3	15.8	21.1	91.3	12.3	2.5	27.1	88.9	83.5	12.3	0.73	0.75	24.7
CG12-627080	7.5	80.5	56.9	75.0	21.5	15.3	20.6	90.5	12.3	2.5	25.5	88.8	59.7	12.8	-1.93	-1.82	17.9
CG12-621017	8.6	72.2	52.8	66.8	21.6	15.6	20.9	90.6	11.3	2.6	27.8	70.2	85.0	13.8	0.87	0.98	15.5
CG12-135018	9.4	75.4	51.4	81.9	20.7	15.2	20.4	90.3	12.3	2.5	28.3	55.8	64.8	13.2	0.89	1.02	19.1
CG12-092002	9.3	84.1	51.4	76.7	20.7	15.2	20.6	89.9	12.3	2.5	28.1	69.7	61.7	13.2	-2.76	-2.66	21.7
CG12-135022	9.7	72.3	29.0	66.7	20.5	14.9	20.1	90.2	11.5	2.4	28.3	95.0	79.5	11.3	1.51	1.47	14.8
CG12-333017	7.5	73.1	47.3	78.6	21.4	15.1	20.7	88.7	12.5	2.5	26.3	95.5	48.1	12.3	2.53	2.65	17.2
CG12-326001	8.0	79.9	43.4	65.6	21.7	16.1	21.7	90.1	11.3	2.4	27.8	90.3	72.4	12.9	-0.41	-0.31	16.5
CG12-135033	9.8	63.1	29.0	71.8	20.5	14.6	19.7	89.9	11.5	2.4	29.6	98.5	92.0	11.3	1.51	1.51	14.6
CG12-030018	6.6	72.9	55.9	80.4	21.0	15.5	21.0	90.1	12.0	2.6	25.4	50.9	72.9	13.8	2.29	2.16	25.3
CG12-335003	7.5	68.7	25.4	71.6	21.6	16.8	22.1	92.3	11.8	2.3	27.3	90.5	80.9	11.6	-0.93	-0.97	25.9
CG12-523028	7.5	77.2	44.4	65.1	20.5	14.9	20.0	90.1	13.3	2.5	26.5	86.8	74.4	13.1	0.14	0.08	14.2
CG12-610018	7.6	80.8	47.3	63.7	20.2	16.2	21.6	91.2	12.5	2.4	27.0	73.9	78.3	13.0	-1.35	-1.31	11.6
CG12-040023	8.1	76.2	46.6	77.3	21.7	15.0	20.6	88.5	13.0	2.3	27.4	86.3	67.9	13.9	1.27	1.25	18.5
CG12-338007	7.4	52.6	50.2	58.2	21.5	15.7	20.9	91.4	11.5	2.4	25.0	89.0	77.3	14.6	-1.02	-1.07	19.2
CG12-324003	5.9	57.1	47.5	79.3	20.5	14.7	20.0	89.3	12.5	2.5	23.7	88.0	78.2	12.8	1.20	1.22	15.5
CG12-626020	7.7	73.6	55.9	66.3	21.7	16.0	21.1	91.9	11.8	2.3	25.4	82.3	87.8	14.0	0.50	0.48	12.0
CP73-1547	7.6	81.8	57.6	79.3	21.2	16.2	21.4	92.3	10.0	2.3	29.5	69.5	84.0	11.8	0.10	-0.03	21.1
CG12-005023	6.5	81.4	55.9	75.3	21.0	16.3	21.6	91.7	12.0	2.4	24.5	58.1	79.6	13.8	-1.39	-1.44	16.1
CG12-333050	6.4	70.1	42.0	76.7	20.8	15.5	21.1	89.8	11.3	2.3	24.4	82.8	68.2	11.9	-0.55	-0.71	21.9
CG12-337006	7.7	63.8	40.1	59.1	21.7	16.2	21.5	91.4	10.8	2.3	26.6	98.8	84.2	13.9	-0.64	-0.74	11.4
CG12-318018	6.7	48.0	58.8	78.8	20.8	16.1	21.5	91.4	12.0	2.2	26.5	90.9	91.3	12.5	-1.17	-1.23	20.2
CG12-032010	8.1	75.6	46.6	72.2	21.7	15.1	20.6	89.0	13.0	2.2	27.6	74.2	66.6	13.9	-0.44	-0.58	23.9
CG12-324191	6.3	78.7	53.8	64.7	21.7	16.6	22.3	90.8	9.8	2.2	25.0	74.7	88.2	13.6	-1.93	-1.98	15.4
CG12-260002	7.8	58.9	58.8	76.9	20.0	15.1	20.4	90.1	12.0	2.2	28.0	85.3	73.7	11.6	-0.02	-0.17	23.1
CG12-324070	5.3	66.4	46.3	77.6	21.5	16.0	21.9	89.1	11.8	2.1	23.0	96.1	93.6	13.2	-2.22	-2.28	15.0

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

Los resultados obtenidos para los dos índices mostraron que las diferencias entre las variedades fueron pequeñas y el orden final de las fracciones superiores, dispuesto discrecionalmente en el 10%, mostraron que las variedades sujetas a selección fueron las mismas para ambos índices (Tabla 4.26), mostrando esencialmente los mismos resultados (los resultados completos sobre índices se presentan en el Anexo 3).

Tabla 4.26. Resumen de las Salidas de RINDSEL (Pérez-Elizalde *et al.*, 2014) para los índices ESIM y RESIM para el conjunto de datos medio_flor_plantía.

Índice	Estadístico	Valor del estadístico	Fracción superior (10% Variedades Seleccionadas)	Valor del índice
ESIM	Covarianza entre el índice y el VMG	0.122	CG12-335003	2.533
	Varianza del índice	0.897	CG12-324191	2.296
	Varianza del VMG	0.012	CG12-146003	2.290
	Correlación entre el índice y el VMG	0.999	CP72-2086 (Testigo)	1.510
RESIM	Covarianza entre el índice y el VMG	0.122	CG12-335003	2.654
	Varianza del índice	0.897	CG12-324191	2.409
	Varianza del VMG	0.012	CG12-146003	2.160
	Correlación entre el índice y el VMG	0.999	CG12-005023	1.511
			CP72-2086 (Testigo)	1.474

VMG: Valor de la Mejora Genética (*Breeding value*).

4.13.1.4 CPs como índices de desempeño para el conjunto de datos medio_flor_plantía

Los Índices RESIM y ESIM se obtuvieron a partir de un ACP. En la sección de Introducción se presentaron los elementos teóricos sobre su cálculo. Las funciones lineales de los ACPs, en sí, también pueden ser utilizadas como funciones que incluyan a todas las variables en estudio y, a su vez, pueden ser utilizadas como índices que resuman el desempeño de los genotipos en estudio. En este otro epígrafe se discutirá los resultados de las funciones desarrolladas en un ACP que, aplicadas a cada observación, generan un resultado (*score*) que puede ser utilizado como una variable de caracterización y, en consecuencia, pueden ser utilizados para la valoración genotípica de los materiales en evaluación.

En el epígrafe 4.10 se presentaron los resultados de las CPs en los conjuntos de datos que se utilizaron como modelos. En el epígrafe 4.10.1 se evaluó lo referente al conjunto de datos **medio_flor_plantía**, de tal manera que las tres primeras CPs explicaron el 53.4% del total de la variabilidad detectada en ese conjunto de datos, para una varianza total en ese conjunto de datos de 14.0. Considerando que la CP1 estuvo asociada a las variables de concentración de azúcar (**pol_6**, **brixjugo_7**, **purity_8** y **brixprom_1**), entonces esta CP1 fue un índice de desempeño relacionado con esas variables. La CP2 estuvo relacionada con variables de acumulación de biomasa (**peso5stalk_2**, **diam_11**, **fiber_14**, principalmente) y, en consecuencia, fue un índice de desempeño asociado a esas variables. Por otro lado, la CP3 también estuvo asociada a variables de acumulación de biomasa y calidad del tallo (**flower_12**, **pith_13**, **alt_10**, **peso5stalk_2**, **oqueleve_3**), por lo que es otro índice asociado a ese tipo de variables (Tabla 4.11 y Figura 4.16). Por tanto, pudo realizarse una valoración genotípica según cada CP y, en consecuencia, obtener un ordenamiento jerárquico de los materiales genéticos en evaluación.

La CP2 presentó el valor más alto de heredabilidad, lo cual resultó consistente con los valores de heredabilidad de las variables que más “carga” presentaron en esta CP, que fueron **peso5stalk_2** y **diam_11**, principalmente. Por otra parte, la CP1 tuvo un relativo bajo valor de heredabilidad, lo cual está relacionado con los valores de heredabilidad de las variables relacionadas con la concentración de azúcar, aparentemente más influidas por el ambiente (varianza de 1.166) (Tabla 4.27).

Tabla 4.27. Resumen de indicadores para los índices CP1, CP2, y CP3.

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY)	Varianza ambiental (RESIDUOS)	H ²	Valor estimado de Mejora genética (VMG) =H ² *(ENTRY + RESIDUOS)
CP1	0.05661	0.844	1.166	0.591	1.188
CP2	-0.02078	1.304	1.034	0.716	1.674
CP2	0.00064	0.887	0.870	0.670	1.177

H²= heredabilidad en el sentido amplio

Se comprobó que no es posible seleccionar la misma variedad en ninguna de las tres CPs, excepto la variedad CP73-1547, la cual es un testigo (Tabla 4.28). Esto indica que ninguna de las variedades, excepto uno de los testigos, cumple las premisas de las tres primeras CPs, ortogonales entre sí. La primera CP sería la primera opción para discriminar a la fracción superior (10%) del resto de los genotipos, ya que acumuló la explicación del 24.3% del total de la variabilidad (Tabla 4.10). Sin embargo, se dejaría fuera cerca del 30% de variabilidad asociada a otros caracteres; esto es, el porcentaje de la varianza de la CP2 más el porcentaje de la varianza de la varianza de la CP3. Esto ocurriría porque no se considerarían las variables asociadas a las CP2 (**peso5stalk_2, diam_11, fiber_14**) y CP3 (**flower_12, pith_13, alt_10, peso5stalk_2, oqueleve_3**). Al seleccionar solamente por la CP1, se considerarían solo variables asociadas a la concentración de sacarosa (**pol_6, brixjugo_7, purity_8** y **brixprom_1**). Una solución a este problema es realizar una selección anidada, según se explicó en la sección de Materiales y Métodos, epígrafe 3.6.6.4.

En el presente caso, se realizó primeramente un Análisis de Varianza solamente para la variable CP1, cuyos datos siguieron una distribución normal (Figura 4.29). Se ordenaron los genotipos según los valores más altos de cada CP para una fracción del 25% (ver sección de Materiales y Métodos epígrafe 3.6.6.4). Se eligió el 25% de la fracción superior para disponer de materiales genéticos suficientes que pudieran ser tomados en cuenta para la selección final, después de incorporar los efectos asociados con la CP2 y la CP3 (Tabla 4.28). Se siguió con la selección de acuerdo con los *scores* de las CP2 y CP3, de tal manera que, al final, pudo alcanzarse la fracción superior del 10%. La elección de la presión de selección del 10% se explicó en la sección de Materiales y Métodos, para todos los índices estudiados.

Tabla 4.28. Ordenación de los genotipos (fracción del 25%) por valores de las CPs 1, 2 y 3 (CP1, CP2, y CP3).

Genotipos según CP1	Score	Genotipos según CP2	Score	Genotipos según CP3	Score
CG12-337006	3.397	CG12-335003	-0.530	CG12-032010	-2.740
CG12-610018	2.464	CG12-253007	2.344	CG12-092002	-1.078
CP72-2086	2.356	CG98-78	0.810	CG12-324191	1.107
CG12-253007	2.344	CP73-1547	1.204	CG12-326001	-0.584
CG12-135018	1.597	CG12-040023	-0.896	CP73-1547	1.204
CG12-621017	1.271	CG12-620011	0.825	CG12-032010	-2.740
CP73-1547	1.204	CG12-030018	0.535	CG12-092002	-1.078

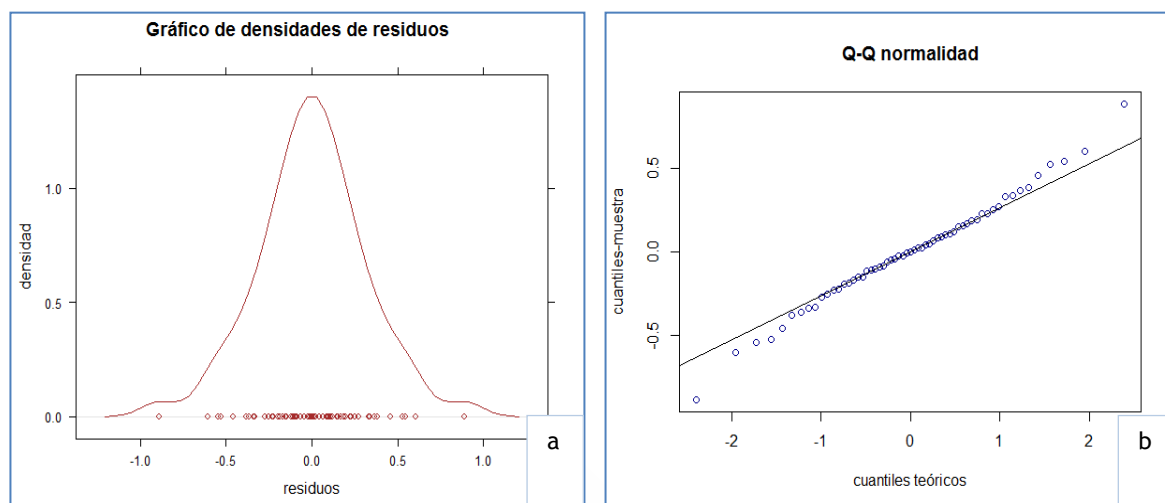


Figura 4.29. Gráficos de densidades (a) y cuantiles (b) para la CP1 del conjunto de datos medio_flor_plantía.

Para un modelo lineal mixto (ver Materiales y Métodos), donde el factor genotipos (ENTRY) se consideró como fijo y el resto aleatorios, se encontró una diferencia significativa entre genotipos ($\text{Pr}>F = 0.025^{**}$). Posteriormente, se realizó una comparación de medias por pares entre genotipos a partir del **ANOVA** (Tabla 4.27) en el que se encontraron diferencias estadísticas.

Una vez hecha la discriminación para la CP1 por el **ANOVA** y la separación de promedios, dentro de la fracción seleccionada por la CP1 (discrecionalmente el 25%), se pudo proceder a seleccionar una fracción de estos genotipos con los mejores valores según los *scores* para la CP2. Seguidamente, dentro de esta última selección, pudieron seleccionarse los genotipos con los más altos valores de la CP3, hasta ajustar la selección del 10% de la fracción superior del total de variedades en evaluación, que para este caso significó encontrar las mejores tres variedades (Tabla 4.29). De esta forma se pudieron considerar las tres CPs en conjunto, de tal manera que la selección se haga dentro de un buen margen del total de la variabilidad explicada para estas tres componentes, es decir, el 53.4% del total de la variabilidad para este conjunto de datos.

Tabla 4.29. Fracción superior del 25% (8 variedades) inicialmente seleccionada por la CP1, seguida de la selección del 10% de la fracción superior, obtenidas por las CP2 y CP3.

Genotipo (ENTRY)	Separación de medias según Tukey para la CP1	CP1	Genotipo (ENTRY)	CP2	Genotipo (ENTRY)	CP3
CG12-337006	DG	3.397	CG12-253007	1.721	CG12-610018	-1.508
CG12-610018	ADFG	2.464	CP73-1547	1.465	CG12-135018	-1.5
CP72-2086 testigo	DFG	2.356	CG12-621017	0.463	CG12-621017	-1.203
CG12-253007	ABCDEFGF	2.344	CG12-135018	0.212	CP72-2086 Testigo	-0.168
CG12-135018	ABCDEFGF	1.597	CG12-324191	0.082		
CG12-621017	D	1.271	CP72-2086 Testigo	0.057		
CP73-1547	ABCDEFGF	1.204				
CG12-324191	ADEFG	1.107				

Los mejores genotipos (Tabla 4.29) según la concentración de sacarosa (CP1, **pol_6**, **brixjugo_7**, **purity_8**, **brixprom_1**; Figura 4.16.a) fueron en total 8 (~25% de la fracción superior): CG12-253007, CP73-1547, CG12-621017, CG12-135018, CG12-324191, CP72-2086, CG12-337006, y CG12-610018. De estos genotipos, según los *scores* correspondientes a la CP2, los que mejor se expresaron con respecto a las variables relacionadas con la acumulación de biomasa (principalmente **peso5stalk_2**, **diam_11**, **fiber_14** y **oqueinter_4**; Figura 4.16.b) fueron: CG12-253007, CP73-1547, CG12-621017, CG12-135018, y CG12-324191. Finalmente, de acuerdo con los *scores* correspondientes en la CP3 (**flower_12**, **pith_13**, **alt_10**, **peso5stalk_2** y **oqueleve_3**; Figura 4.16), los genotipos con el mejor desempeño fueron CG12-610018, CG12-135018, y CG12-621017. En este caso, debido a que las cargas (*Eigenvectors*) fueron de orden negativo (Tabla 4.11), los genotipos con mejor desempeño fueron aquellos cuyos *scores* eran los más bajos (10% de la fracción superior de todos los genotipos en evaluación). De esta manera, se seleccionaron a los genotipos que, equilibradamente, presenten el mejor desempeño fenotípico.

Con el anterior procedimiento se buscó un equilibrio en la selección entre concentración de sacarosa, acumulación de biomasa y calidad interna de los tallos. A primera vista, al observar los listados de genotipos, podría esperarse que el genotipo CG12-337006 fuera el primer seleccionado según la CP1. Sin embargo, para el resto de caracteres relacionados con la acumulación de biomasa y la calidad interna de los tallos, este genotipo no cumple

satisfactoriamente con todos los caracteres, por tanto no fue seleccionado. Lo que se buscó con este sistema de selección anidada fue la selección de aquellos genotipos que presentaran un equilibrio entre concentración de sacarosa, acumulación de biomasa y calidad interna de los tallos.

4.13.1.5 TAH (t De Azúcar Por ha) y BLUP (*Best Linear Unbiased Prediction*)

Originalmente, TAH no se consideró como un índice. Sin embargo, por la forma del cálculo, este valor puede ser considerado como tal. En la sección de Materiales y Métodos se indicaron los detalles sobre el cálculo de TAH y algunos elementos sobre las bases teóricas de éste carácter. En todo caso, TAH es un valor importante de referencia por cuanto es la variable de mayor importancia en la selección de las nuevas variedades de caña de azúcar en el programa de mejora genética de caña de azúcar de CENGICANA.

En un resumen de los valores relacionados con el índice de desempeño TAH (Tabla 4.30), se aprecia el valor fenotípico promedio (en este caso, también reflejado en la varianza genotípica); el valor de la heredabilidad (H^2) relativamente bajo para este carácter en este conjunto de datos; y el valor estimado de mejoramiento genético (*Estimated Breeding Value*), que es el valor que equivale a la varianza específica relacionada con el desempeño fenotípico imputable a los efectos de los genes relacionados con el carácter.

Tabla 4.30. Resumen de indicadores del índice TAH (t de azúcar por ha).

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY): variedad	Varianza ambiental (RESIDUOS)	H^2	Valor estimado de Mejora genética = $H^2 \times (\text{ENTRY} + \text{RESIDUOS})$
TAH	18.76 ± 5.71	5.980	25.610	0.318	10.055

H^2 = heredabilidad en el sentido amplio

El valor fenotípico medio fue bajo (Tabla 4.30) comparado con el resto de conjuntos de datos evaluados (Tabla 4.2), solamente superior al conjunto de datos **medio_flor_1ª_soca**, y con un desempeño similar al conjunto **medio_no_flor_1ª_soca**. En general, como se indicó, en este conjunto de datos se recabaron valores bajos para la acumulación de biomasa, como consecuencia de un crecimiento limitado de las plantas debido a las condiciones limitantes del suelo en el que se instaló este experimento. Este conjunto de datos presentó el valor más bajo

de heredabilidad, lo cual se debió a una mayor variabilidad de la respuesta de cada genotipo en cada repetición, lo que se vio también reflejado en un bajo valor de mejora genética (*Breeding value*). En todo caso, los valores de TAH se presentaron como una opción para la valoración de los genotipos en este conjunto de datos.

Por otro lado, el solo uso del TAH como un índice de desempeño pudo ser utilizado en un esquema de análisis para un modelo lineal mixto, donde se puede considerar como efecto fijo a los materiales genéticos (ENTRY), y como efectos aleatorios a los componentes del modelo relacionados con las repeticiones y los bloques incompletos (ver modelo en la sección de Materiales y Métodos).

Para este conjunto de datos se observó que la variable TAH mostró una aceptable distribución de los residuos (Figura 4.30.a), y una tendencia aceptable como variable de distribución normal (Figura 4.30.b); por tanto, los datos, en términos generales, fueron adecuados para su uso como un índice y para ser utilizados en un **ANOVA**.

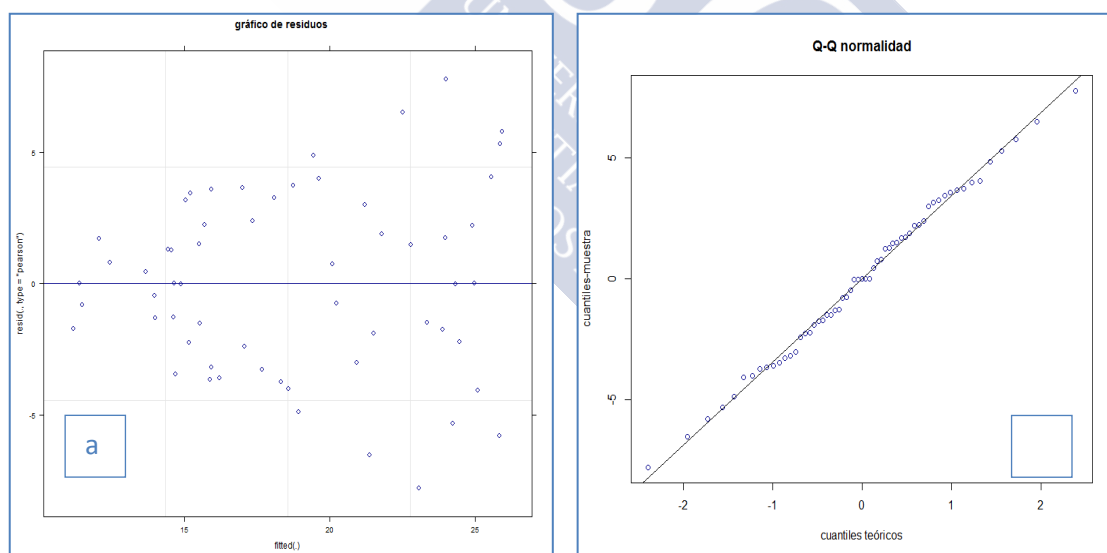


Figura 4.30. Gráficos de residuos (a) y de cuantiles (b) sobre la distribución de los datos de las observaciones para la variable/índice TAH (t de azúcar por ha) en el conjunto de datos “litoral_flor_1ª_soca”.

Con este conjunto de datos se realizó un Análisis de Varianza para un modelo mixto según el modelo lineal aditivo mixto (ver Materiales y Métodos). El factor variedad (ENTRY) se estableció como un efecto fijo, mientras que el resto de factores fueron considerados

aleatorios. La fuente variedad, ENTRY (efecto fijo), no mostró diferencias significativas en los promedios del factor genotipo ($F= 1.4771$; $p>F= 0.175$) para este conjunto de datos.

En el presente estudio, BLUP se aplicó a la variable/índice TAH, ya que refleja en buena medida las tendencias del desempeño de cada genotipo, e incluye una serie de mediciones que involucran medidas de producción de biomasa y concentración de azúcar en cada genotipo en estudio. Por otro lado, el TAH es un indicativo para el fitomejorador de un buen nivel de desempeño de los genotipos en evaluación. BLUP tiene la propiedad de acercar las lecturas hacia la media de la variable (Piepho *et al.*, 2008; Hill y Rosenberg, 1985), y resulta en un valor que reduce el error de sesgo de la variable de la cual es originado, en este caso el TAH.

La estimación de BLUP para TAH (Figura 4.31) mostró una distribución del número de genotipos (frecuencias) según rangos de los valores de BLUP (los cuales van desde -3 hasta +3 en rangos de uno en uno). Se observó que el mayor número de genotipos presentaron valores negativos, aquellos que se asociaron principalmente con genotipos con bajos niveles de desempeño para el TAH.

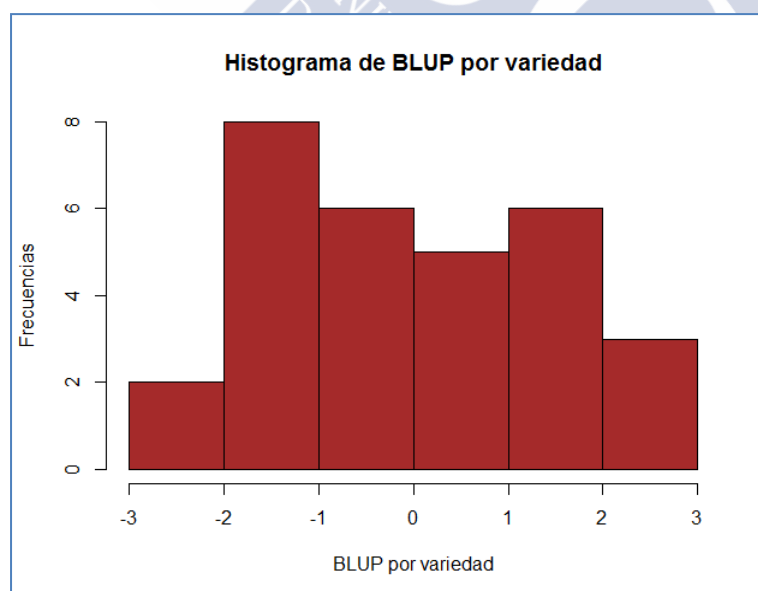


Figura 4.31. Distribución de las frecuencias de los BLUPs.

Por otro lado, y en concordancia con la asociación de cada BLUP con su respectivo valor de TAH (Figura 4.32), una de las propiedades de BLUP es la aproximación de los

estimadores hacia la media de la variable de la cual se deriva (Piepho *et al.*, 2008; Hill y Rosenberg, 1985), de tal manera que pueda utilizarse el estimador como una mejor opción de la valoración genotípica, en este caso con respecto al TAH, ya que BLUP presentará una menor desviación que la que presenta la variable de la cual es estimado.

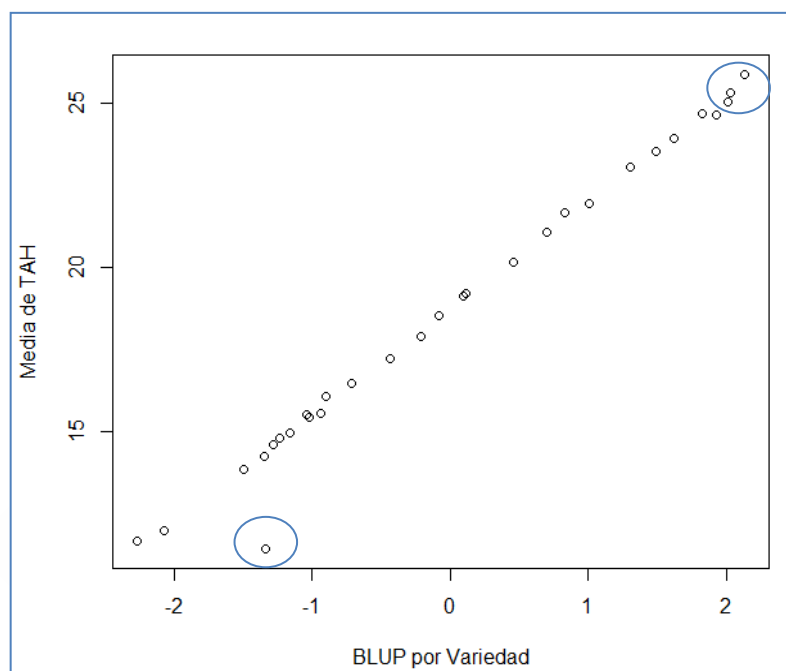


Figura 4.32. Relación BLUP con las medias de TAH.

Los valores que se desviaron de la línea media diagonal en la Figura 4.32 (la media del TAH), son acercados por BLUP a dicha línea, de tal manera que su valoración se acerca más a los valores que son más comunes para todos los genotipos en evaluación. Esta ventaja permite hacer una selección más precisa, utilizando los valores de BLUP en lugar de los valores originales de TAH. En este caso, al aplicar el 10% de presión de selección sobre la base de los BLUPs, se observó que si se hubiese hecho la selección de la fracción superior (10%) por TAH (Tabla 4.31), el mejor tercer genotipo seleccionado hubiera sido CG12-620011, ya que presentó un valor mayor; sin embargo, la estimación insesgada que se hace con el cálculo de BLUP, permite acercar el estimador con más precisión a la media y, en consecuencia, se hace un ajuste que hace que el genotipo sujeto de selección sea el CG12-253007. No se considera para la selección el genotipo CG98-78, ya que este es un testigo dentro de la evaluación. Si bien, en este caso, las diferencias entre los genotipos en cuestión fueron relativamente

pequeñas con respecto al TAH, el uso de BLUP puede ser un elemento relevante para tomar la decisión de seleccionar uno u otro genotipo.

Tabla 4.31. Ordenamiento descendente de la fracción superior (23%) de genotipos según los valores de BLUP.

Genotipo	TAH	BLUP
CG12-335003	25.89	2.131
CG12-030018	25.32	2.025
CG98-78 (Testigo)	25.05	2.013
CG12-253007	24.65	1.929
CG12-620011	24.69	1.821
CG12-032010	23.93	1.615
CP72-2086	23.54	1.492

4.13.1.6 Regresión Logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos medio_flor_plantía

Con los resultados obtenidos hasta ahora, en este conjunto de datos y, así mismo, con el resto de conjunto de datos, resulta en cierta manera complicado establecer cuál o cuáles de los índices pueden representar una mejor opción para su uso en cada conjunto de datos. La competencia de selección por parte del fitomejorador resulta, en buena medida, un buen nivel de comparación si se cumple con algunas condiciones por parte del mismo: experiencia, preparación académica, conocimiento claro y muy específico del cultivo, éxito en la liberación de genotipos, etcétera.

Para el presente estudio se ha definido que la selección por competencia cumple en buena medida con las condiciones indicadas. En ese sentido, se procedió a realizar una comparación del conjunto de índices para cada genotipo en evaluación con respecto a la evaluación y selección por competencia por parte del fitomejorador.

Debido a que el resultado final de una selección por parte del fitomejorador es positiva (si) o negativa (no), ésta se constituye como una variable binomial con solo dos posibles resultados. Por tal razón, se procedió a realizar una serie de ajustes de modelos lineales Logit (Manning, 2007; Everitt, 2014) con el objeto de establecer la calidad de dichas relaciones. El

procedimiento se presentó con mayor detalle en la sección de Materiales y Métodos y la sección de Introducción.

En general, según el **Pseudo- r^2** ningún índice mostró significación con la base de comparación (Tabla 4.32), por lo que no se mostraron como una opción aceptable en este conjunto de datos. Si bien se encontró esta falta de relación, es importante señalar que solamente en este conjunto de datos se encontró tal tendencia. En el resto de conjuntos de datos se obtuvo relaciones con valores que van de medios a altos. Se observa en la columna de Eficiencia que los dos índices de más alta eficiencia fueron TAH y BLUP, con los valores más altos de Pseudo- r^2 (Tabla 4.32).

Tabla 4.32. Valores de Pseudo r^2 de McFadden de una regresión Logit para el modelo general $\ln[p/(1-p)] = a + bX + e$ para el conjunto de datos medio_flor_plantía.

Índice	Modelo	Prueba de significación $Pr(>\chi^2)$	Pseudo- r^2 con la base <i>selección por competencia</i>	Eficiencia con respecto al r^2 más alto
TAH	$\ln[Si/(1-Si)] = -2.59561 + 0.1473TAH$	0.088	0.070	1.00
BLUP	$\ln[Si/(1-Si)] = 0.1492 + 0.4599BLUP$	0.109	0.062	0.88
CP2	$\ln[Si/(1-Si)] = 0.1486 + 0.3479CP2$	0.229	0.035	0.50
ESIM	$\ln[Si/(1-Si)] = 0.1371 + 0.2331ESIM$	0.376	0.019	0.28
RESIM	$\ln[Si/(1-Si)] = 0.1371 + 0.2248RESIM$	0.383	0.018	0.26
CP3	$\ln[Si/(1-Si)] = 0.1370 + 0.2853CP3$	0.386	0.018	0.26
CP1	$\ln[Si/(1-Si)] = 0.1406 - 0.1068CP1$	0.643	0.005	0.07
KN	$\ln[Si/(1-Si)] = 0.13353 - 0.01292KN$	0.981	<0.0001	0.00
Smith	$\ln[Si/(1-Si)] = 0.1335 - 0.0049Smith$	0.994	<0.0001	0.00

a=intercepto, b=coeficiente, X=Índice, e=término del error, p=probabilidad de ocurrencia de "sí", 1-p=probabilidad de ocurrencia de "no".

Es razonable que tanto los valores de TAH como de BLUP fueran los que presentaron los índices de correlación más altos con la base de selección por competencia, ya que la ***selección por competencia*** está altamente influida por dos grandes factores: producción de biomasa y concentración de azúcar. En este sentido TAH fue calculado a partir de la acumulación de biomasa y de la concentración de azúcar, y BLUP fue el mejor estimador insesgado de TAH.

Si se omite el uso de TAH y BLUP de TAH en este conjunto de datos, el índice CP2 se presenta como uno de los mejores, con una eficiencia del 50%. El índice CP2 está asociado con algunas variables de acumulación de biomasa (**peso5stalk_2** y **diam_11**). En todo caso es importante remarcar que ninguno de los índices presentó un buen nivel de asociación con la base de comparación en la selección (***selección por competencia***), por lo que en este caso no

son aplicables en la forma que fueron calculados. En este sentido, puede hacerse una re-selección de las variables que se utilizan para el cálculo de los índices y verificar las tendencias, es decir, no utilizar todas las 14 variables y, de esta manera, verificar si puede encontrar un mejor ajuste de los índices.

Los valores de $Pr(>\chi^2)$ para los modelos de TAH y BLUP fueron no significativos ($p > 0.05$), los valores de Pseudo- r^2 de McFadden también fueron bajos (Tabla 4.32), como lo fueron los niveles de regresión de las curvas correspondientes (Figura 4.33). Por otro lado, los índices Smith y KN mostraron una relación con la base de comparación prácticamente nula.



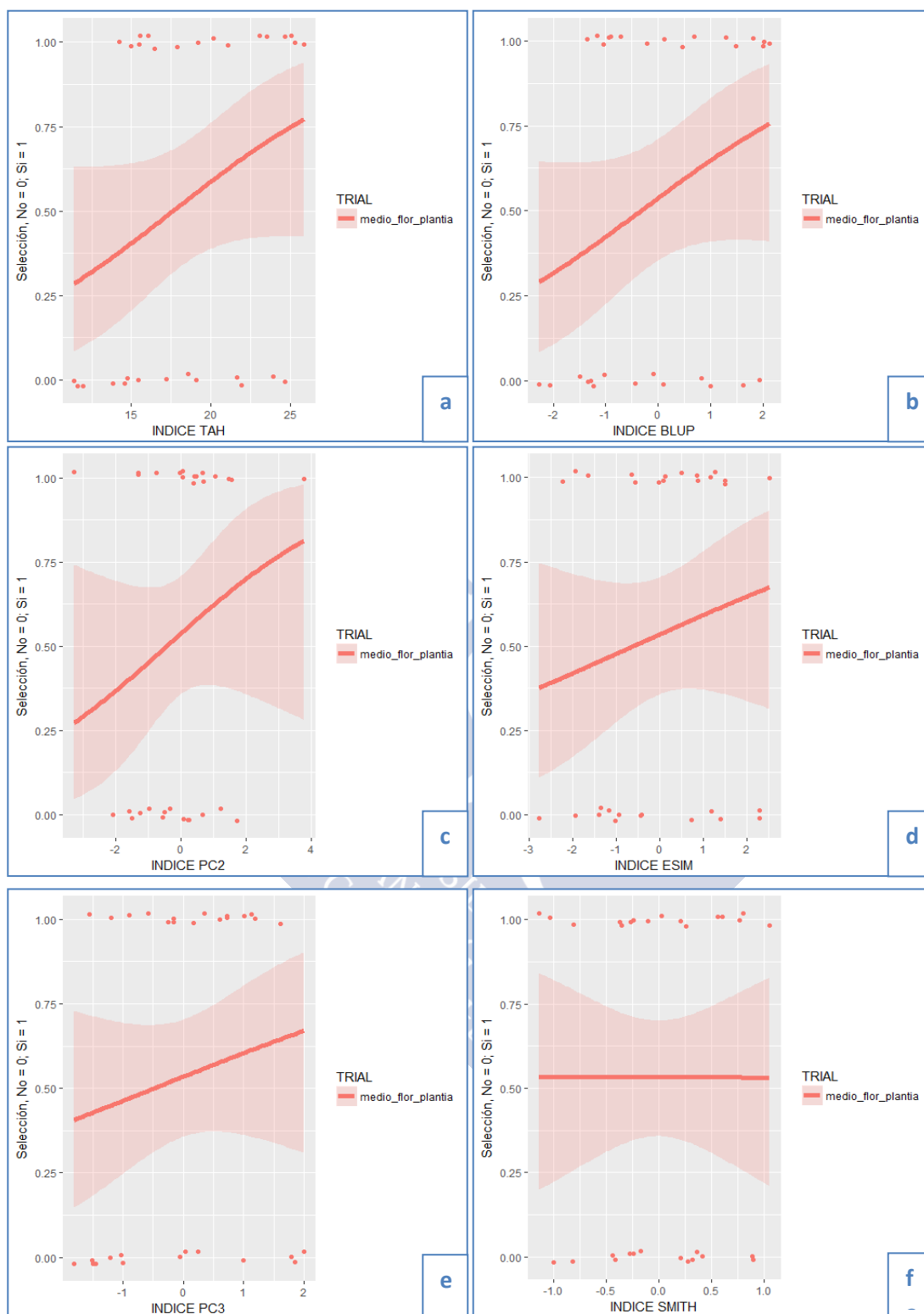


Figura 4.33. Gráficos de Regresión lineal Logit para la relación entre la base de comparación de selección por competencia (“Sí” o “No”) con seis índices de selección: a) TAH (t de azúcar por ha); b) BLUP de TAH; c) CP2; d) ESIM; e) CP3; y f) Smith.

4.13.2 Índices de selección para el conjunto de datos medio no flor 1ª soca (medio_no_flor_1ª_soca)

4.13.2.1 Análisis de correlación entre índices

Para este conjunto de datos se obtuvo un buen nivel de expresión fenotípica reflejado en elevados valores de TAH, como producto de una mayor acumulación de biomasa en términos del peso de cinco tallos, mayor altura y diámetro. Una particularidad de este experimento (estrato medio, variedades **no_flor**), el cual aglutinó dos conjuntos de datos (plantía y primera soca), es que fue el experimento donde más variedades fueron seleccionadas, siguiendo la base de comparación de este estudio *Selección por competencia del fitomejorador*, debido al elevado nivel de expresión fenotípica que presentaron los genotipos en los dos ciclos de evaluación.

En este otro grupo de correlaciones (Tabla 4.33) se observó el mismo patrón que el que se discutió anteriormente para los índices presentados en el conjunto de datos anterior (Tabla 4.22). Estas asociaciones fueron más marcadas entre el TAH y el BLUP, así como la correlación entre ESIM y RESIM. En este sentido, cada par de índices pueden considerarse como intercambiables o bien considerarse en esencia lo mismo. Las correlaciones entre la CP1 y los índices Smith, RESIM y ESIM indicaron que las variables relacionadas con el azúcar fueron las que influyeron estos índices. Además, se observó que los índices Smith, RESIM y ESIM también fueron afectados por las variables relacionadas con la calidad interna del tallo (**pith_13**, **flower_12** y **oqueleve_3**) y por la variable de acumulación de biomasa **peso5stalk_2**. Lo anterior se desprende de la correlación de estos índices con la CP2. Si bien el TAH recoge en su valor la participación **pol_6** y **peso5stalk_2**, tanto este índice como BLUP no se correlacionaron con ninguno del resto de los índices.

Tabla 4.33. Correlación de Pearson y nivel de significación entre los valores de los nueve índices obtenidos para el conjunto de datos “medio_no_flor_1^a_soca”.

	TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3
TAH		****	NS	NS	NS	NS	NS	NS	****
BLUP	0.99		NS	NS	NS	NS	NS	NS	****
ESIM	0.05	0.07		****	NS	*	****	NS	NS
RESIM	0.10	0.12	0.99		NS	**	****	NS	NS
KN	0.21	0.20	0.00	0.07		NS	NS	NS	**
Smith	0.06	0.04	-0.34	-0.37	0.06		NS	NS	NS
CP1	-0.01	-0.01	-0.54	-0.54	-0.01	0.22		NS	NS
CP2	-0.08	-0.10	-0.21	-0.23	0.04	0.21	0.03		NS
CP3	0.69	0.70	0.10	0.16	0.41	-0.07	0.00	-0.03	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

4.13.2.2 Índice “Smith” y “Restrictive Kempthorne and Nordskog Selection Index” (Índice KN)

En la Tabla 4.34 se presentan los valores promedio de las variables de respuesta y los valores de los índices Smith (en orden descendente) y KN y TAH, éste último como referencia de productividad, para este conjunto de datos.

En general, los índices Smith y KN (salidas de RINDSEL y la selección del 10%) no se correlacionaron ($r = -0.18$, NS), si bien se considera que KN es una derivación teórica de Smith. Posiblemente, debido a las restricciones en las variables en KN (ver Materiales y Métodos epígrafe 3.6.6.2), la correlación entre el índice y el VMG fue menor que el que presentó Smith. Las selecciones que se obtuvieron con KN fueron variedades con valores bajos de **peso5stalk_2** principalmente, es decir, que no ponderó apropiadamente esta variable. En el caso de Smith, sí ponderó apropiadamente el **peso5stalk_2**, pero no la concentración de azúcar (**pol_6**). En este sentido, ambos índices no seleccionaron variedades que combinaran apropiadamente estos dos componentes. Esto está correlacionado con los niveles de heredabilidad de las variables en estudio (Tabla 4.7) de los materiales **no_flor** del estrato medio, en donde las variables que más ponderaron fueron **pith_13**, **fiber_14**, y **diam_11**.

Tabla 4.34. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_no_flor_1^a_soca.

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flowe_12	pith_13	fiber_14	KN	Smith	TAH
CG12-132	9.9	85.7	58.8	79.5	20.4	15.8	21.5	91.8	11.8	2.4	30.0	100.0	100.0	11.23	0.30	0.70	23.5
CP72-2086	6.7	67.7	56.5	79.5	22.9	16.0	21.7	90.9	10.4	2.5	26.8	100.0	100.0	12.57	0.46	0.44	17.0
CG12-008054	8.2	76.3	58.8	79.6	21.5	16.0	21.6	90.2	10.9	2.5	28.4	100.0	100.0	12.36	0.39	0.41	21.7
CG12-133	8.7	77.2	54.6	79.5	21.7	15.7	21.4	91.9	11.7	2.5	27.9	100.0	94.3	11.41	0.60	0.38	19.1
CG12-637001	7.0	74.8	52.9	78.6	21.4	15.6	21.1	91.1	11.0	2.6	28.0	100.0	95.3	11.86	0.72	0.38	12.9
CG12-531020	7.1	83.8	53.0	78.8	21.2	16.7	22.1	91.9	12.0	2.5	23.2	100.0	90.8	12.01	-0.06	0.35	16.7
CG12-317043	6.6	66.6	34.5	79.2	21.5	15.8	21.2	90.4	10.6	2.3	26.6	100.0	100.0	12.12	-0.46	0.35	15.6
CG12-606102	6.0	85.9	52.7	75.8	20.4	15.3	21.0	88.6	15.3	2.5	25.8	84.8	71.0	13.30	0.29	0.31	19.8
CG12-606100	5.8	82.2	58.8	79.2	21.7	16.1	21.6	90.7	12.4	2.4	24.3	100.0	99.0	12.07	-0.11	0.27	17.8
CG12-533002	6.7	86.5	58.8	79.4	21.4	16.2	21.8	90.7	13.6	2.5	24.5	100.0	100.0	11.28	0.03	0.24	22.2
CG12-125	6.5	83.8	57.5	79.0	21.8	16.5	21.9	91.3	12.5	2.4	23.3	100.0	100.0	12.40	-0.38	0.21	19.5
CG12-629011	6.7	67.9	42.4	78.9	21.7	16.3	21.7	91.2	12.7	2.4	24.6	99.7	99.3	13.83	-0.37	0.20	18.6
CG12-627037	5.5	86.5	58.8	79.3	21.2	16.3	21.3	92.1	11.0	2.6	21.2	100.0	100.0	13.94	0.20	0.20	23.3
CG12-324185	7.5	77.0	43.1	79.8	20.9	16.2	21.3	92.0	11.3	2.4	26.8	100.0	96.8	13.62	-0.31	0.17	18.2
CG12-564013	7.2	74.6	58.4	77.8	21.5	15.9	21.3	90.4	11.3	2.4	27.9	84.5	97.5	11.77	-0.04	0.17	13.8
CG12-548002	7.5	79.2	41.5	73.3	20.8	16.2	21.6	91.3	11.4	2.2	24.1	100.0	99.8	11.01	-0.87	0.15	19.9
CG12-332015	6.7	82.3	58.6	79.3	23.0	16.7	22.1	92.1	10.4	2.3	24.2	100.0	100.0	12.62	-0.53	0.13	16.8
CG12-626007	6.3	74.5	47.6	76.9	21.6	15.4	21.0	89.0	13.9	2.3	23.9	84.8	71.0	13.13	-0.35	0.12	15.9
CG12-317020	8.5	67.1	44.3	76.6	21.2	15.9	21.6	90.1	12.7	2.4	27.1	100.0	98.8	10.70	-0.10	0.11	15.7
CG12-324187	6.7	74.1	56.4	79.2	21.1	16.7	21.8	92.9	12.7	2.5	24.1	100.0	99.8	13.17	-0.33	0.09	20.3
CG12-126	8.5	70.5	54.5	78.0	21.1	16.4	21.6	92.0	11.0	2.5	24.0	100.0	100.0	12.25	-0.06	0.05	20.4
CP88-1165	9.2	76.7	46.6	77.7	22.3	16.8	22.2	91.9	10.2	2.5	26.1	84.5	97.5	12.48	-0.12	0.01	17.9
CG12-512013	8.0	76.2	50.2	76.4	22.3	17.4	22.7	93.0	11.0	2.5	25.7	100.0	85.0	12.63	-0.03	0.00	20.3
CG12-135	7.1	73.6	53.2	78.7	21.2	15.9	21.3	90.9	13.5	2.7	26.3	100.0	93.3	12.58	0.40	-0.04	21.3
CG12-324156	7.2	78.7	44.0	79.3	21.9	16.4	21.8	91.6	12.8	2.6	25.9	100.0	66.5	13.39	0.38	-0.06	18.4
CG12-122	7.0	70.6	54.1	79.3	20.9	15.5	20.6	91.2	12.6	2.5	23.6	100.0	100.0	14.62	0.11	-0.06	20.5
CG12-318012	7.4	75.9	33.4	79.0	22.0	16.6	21.9	92.4	11.2	2.7	29.6	100.0	94.8	13.68	0.24	-0.08	22.1
CG12-121	6.2	76.5	54.1	79.2	22.1	15.7	21.0	90.6	12.2	2.6	26.3	100.0	97.3	12.71	0.42	-0.08	15.4
CG12-534006	7.4	83.8	58.4	79.3	21.6	16.0	21.3	91.2	12.3	2.3	25.1	100.0	100.0	10.82	-0.47	-0.09	16.0
CG12-333077	6.3	80.1	57.2	76.9	22.9	17.0	22.4	92.3	12.6	2.6	24.4	100.0	93.5	12.39	0.12	-0.10	18.2
CG12-512009	7.6	73.9	53.2	77.6	22.7	16.7	22.0	92.1	11.2	2.6	25.6	100.0	94.3	12.68	0.16	-0.11	19.2
CG12-318022	8.3	72.2	46.2	72.7	21.9	16.6	22.0	92.2	10.5	2.6	27.2	100.0	94.8	13.33	0.04	-0.12	16.3
CG12-130	9.9	70.4	55.8	78.3	20.2	15.9	21.0	91.6	11.0	2.5	28.1	100.0	100.0	11.52	-0.21	-0.13	24.5
CG12-324147	6.2	83.9	54.7	78.0	21.9	16.1	21.6	90.9	14.9	2.5	23.1	100.0	65.5	13.57	0.08	-0.13	20.8
CG12-360009	7.9	66.6	50.3	79.3	22.0	16.4	21.6	92.1	10.6	2.6	26.0	100.0	100.0	13.15	0.21	-0.14	18.1
CG12-034098	8.4	78.2	59.0	79.0	22.7	16.9	22.3	91.7	11.7	2.6	26.5	100.0	100.0	12.35	0.18	-0.17	19.8
CG12-229005	8.6	78.4	48.7	76.1	21.2	16.3	21.7	91.1	14.3	2.7	27.7	100.0	97.8	12.36	0.45	-0.18	31.2
CG12-324084	7.3	76.4	49.1	72.7	21.9	16.0	21.3	90.8	12.0	2.3	24.3	100.0	96.0	13.21	-0.52	-0.22	19.4

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flowe_12	pith_13	fiber_14	KN	Smith	TAH
Continuación Tabla 4.34																	
CG12-606012	5.5	75.2	46.2	79.4	23.3	17.0	22.2	92.9	11.4	2.3	23.2	100.0	94.5	11.92	-0.63	-0.24	13.0
CG12-332013	7.4	84.4	58.8	79.4	22.9	16.7	21.9	92.8	10.6	2.3	24.0	100.0	100.0	12.74	-0.74	-0.24	14.5
CG12-317044	7.1	65.9	55.3	78.8	22.5	16.6	21.8	92.4	11.0	2.6	25.9	100.0	95.3	12.19	0.14	-0.26	17.0
CG12-570001	6.3	74.6	29.3	72.9	21.4	16.2	21.3	92.6	12.1	2.3	23.3	100.0	95.3	11.92	-0.73	-0.28	14.3
CG12-124	6.0	80.3	58.8	79.2	23.5	16.5	21.9	91.7	13.0	2.5	24.1	100.0	100.0	13.39	-0.01	-0.28	16.9
CG12-057019	7.6	81.1	59.0	79.3	23.0	16.5	21.9	91.2	11.5	2.8	28.5	100.0	97.3	11.19	0.72	-0.28	17.0
CG12-330036	7.2	77.1	57.3	79.2	22.3	17.1	22.2	93.7	12.5	2.5	24.2	100.0	100.0	12.43	-0.38	-0.32	20.4
CG98-78	9.8	75.7	38.5	79.1	22.2	16.3	21.5	92.3	10.0	2.7	28.1	100.0	92.5	13.06	0.43	-0.43	16.1
CG12-337009	7.3	67.3	51.6	77.8	22.3	16.4	21.5	92.5	11.1	2.7	26.1	100.0	93.5	12.84	0.31	-0.44	16.3
CG12-MSING#4	7.4	73.9	39.0	79.1	22.2	16.4	21.5	92.7	12.2	2.6	26.0	99.7	91.8	13.86	-0.02	-0.48	23.5
CG12-123	7.3	76.4	58.8	79.5	22.2	15.9	21.0	91.9	12.1	2.7	22.7	100.0	100.0	14.20	0.42	-0.48	17.3

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueveve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flowe_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

Análogamente a lo que se observó en el conjunto de datos **medio_flor_plantía**, para el ensayo **medio_no_flor_1ª_soca** las selecciones que se realizaron en Smith y KN fueron diferentes excepto por la variedad **CG12-637001**, que fue la última opción para Smith y la primera para KN (Tabla 4.35).

Tabla 4.35. Resumen de las salidas de RINDSEL para los índices “Smith” y “KN” para el ensayo medio_no_flor_1ª_soca.

Índice	Estadístico	Valor del estadístico	Fracción superior (10% variedades seleccionadas)	Valor del índice
Smith	Covarianza entre el índice y el VMG	6.724	CG12-132	0.696
	Varianza del índice	0.110	CP72-2086	0.444
	Varianza del VMG	189.949	CG12-008054	0.413
	Correlación entre el índice y el VMG	0.999	CG12-133	0.381
			CG12-637001	0.380
KN	Covarianza entre el índice y el VMG	1.648	CG12-637001	0.722
	Varianza del índice	0.022	CG12-057019	0.719
	Varianza del VMG	189.949	CG12-133	0.595
	Correlación entre el índice y el VMG	0.810	CP72-2086	0.463
			CG12-229005	0.451

4.13.2.3 *Eigenanalysis-based Selection Index Method (ESIM) y Restrictive Eigenanalysis-based Selection Index Method (RESIM)*

Los valores de los índices ESIM y RESIM para el conjunto de datos **medio_no_flor_1ª_soca** se presentan en la Tabla 4.36, con los valores promedio de cada variable y el de los dos índices mencionados, así como los valores para TAH como términos de referencia.

Como se explicó en la sección de correlaciones, se comprobó que ambos índices, ESIM y RESIM (salidas de Rindsel), fueron prácticamente intercambiables (Tabla 4.37) y, en general, presentaron las mismas selecciones, excepto por los genotipos CG12-606012 y CG12-057019, los cuales fueron las últimas opciones de selección para **ESIM** y **RESIM**, respectivamente. Ambos índices privilegiaron la concentración de azúcar (**pol_6**) sobre el resto de variables. Sin embargo, ambos índices no fueron capaces de recoger apropiadamente la participación de las variables asociadas a la acumulación de biomasa, la cual es también muy importante. En este sentido, estos dos índices parecen no equilibrar las cargas correspondientes para los dos tipos de variables: acumulación de biomasa (**peso5stalk_2**, **alt_9**, **diam_11**) y la concentración de sacarosa (**pol_6**, **brix_7** y **brixprom_1**).

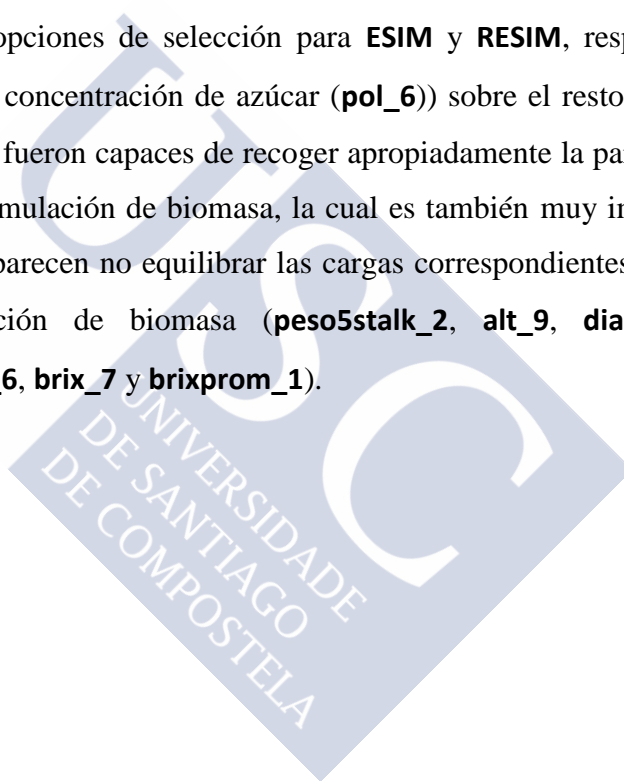


Tabla 4.36. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo medio_no_flor_1ª_soca

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flowe_12	pith_13	fiber_14	ESIM	RESIM	TAH
CG12-229005	8.6	78.4	48.7	76.1	21.2	16.3	21.7	91.1	14.3	2.7	27.7	100	98	12.4	0.58	0.78	31.3
CG12-130	9.9	70.4	55.8	78.3	20.2	15.9	21.0	91.6	11.0	2.5	28.1	100	100	11.5	-1.71	-1.41	24.5
CG12-132	9.9	85.7	58.8	79.5	20.4	15.8	21.5	91.8	11.8	2.4	30.0	100	100	11.2	-1.11	-1.03	23.5
CG12-MSING#4	7.4	73.9	39.0	79.1	22.2	16.4	21.5	92.7	12.2	2.6	26.0	100	92	13.9	0.41	0.58	23.5
CG12-627037	5.5	86.5	58.8	79.3	21.2	16.3	21.3	92.1	11.0	2.6	21.2	100	100	13.9	-0.70	-0.33	23.3
CG12-533002	6.7	86.5	58.8	79.4	21.4	16.2	21.8	90.7	13.6	2.5	24.5	100	100	11.3	-0.05	-0.07	22.3
CG12-318012	7.4	75.9	33.4	79.0	22.0	16.6	21.9	92.4	11.2	2.7	29.6	100	95	13.7	1.33	1.53	22.1
CG12-008054	8.2	76.3	58.8	79.6	21.5	16.0	21.6	90.2	10.9	2.5	28.4	100	100	12.4	-0.61	-0.60	21.7
CG12-135	7.1	73.6	53.2	78.7	21.2	15.9	21.3	90.9	13.5	2.7	26.3	100	93	12.6	-0.94	-0.65	21.3
CG12-324147	6.2	83.9	54.7	78.0	21.9	16.1	21.6	90.9	14.9	2.5	23.1	100	66	13.6	-0.19	-0.22	20.8
CG12-122	7.0	70.6	54.1	79.3	20.9	15.5	20.6	91.2	12.6	2.5	23.6	100	100	14.6	-2.81	-2.48	20.5
CG12-330036	7.2	77.1	57.3	79.2	22.3	17.1	22.2	93.7	12.5	2.5	24.2	100	100	12.4	2.36	2.33	20.4
CG12-126	8.5	70.5	54.5	78.0	21.1	16.4	21.6	92.0	11.0	2.5	24.0	100	100	12.3	-0.01	0.10	20.4
CG12-512013	8.0	76.2	50.2	76.4	22.3	17.4	22.7	93.0	11.0	2.5	25.7	100	85	12.6	3.67	3.53	20.3
CG12-324187	6.7	74.1	56.4	79.2	21.1	16.7	21.8	92.9	12.7	2.5	24.1	100	100	13.2	0.75	0.90	20.3
CG12-548002	7.5	79.2	41.5	73.3	20.8	16.2	21.6	91.3	11.4	2.2	24.1	100	100	11.0	-0.97	-1.20	19.9
CG12-034098	8.4	78.2	59.0	79.0	22.7	16.9	22.3	91.7	11.7	2.6	26.5	100	100	12.3	2.43	2.27	19.8
CG12-606102	6.0	85.9	52.7	75.8	20.4	15.3	21.0	88.6	15.3	2.5	25.8	85	71	13.3	-2.61	-2.46	19.8
CG12-125	6.5	83.8	57.5	79.0	21.8	16.5	21.9	91.3	12.5	2.4	23.3	100	100	12.4	0.40	0.21	19.5
CG12-324084	7.3	76.4	49.1	72.7	21.9	16.0	21.3	90.8	12.0	2.3	24.3	100	96	13.2	-1.16	-1.40	19.4
CG12-512009	7.6	73.9	53.2	77.6	22.7	16.7	22.0	92.1	11.2	2.6	25.6	100	94	12.7	1.52	1.41	19.2
CG12-133	8.7	77.2	54.6	79.5	21.7	15.7	21.4	91.9	11.7	2.5	27.9	100	94	11.4	-1.04	-1.02	19.1
CG12-629011	6.7	67.9	42.4	78.9	21.7	16.3	21.7	91.2	12.7	2.4	24.6	100	99	13.8	-0.11	-0.28	18.6
CG12-324156	7.2	78.7	44.0	79.3	21.9	16.4	21.8	91.6	12.8	2.6	25.9	100	67	13.4	0.66	0.79	18.4
CG12-333077	6.3	80.1	57.2	76.9	22.9	17.0	22.4	92.3	12.6	2.6	24.4	100	94	12.4	2.66	2.49	18.2
CG12-324185	7.5	77.0	43.1	79.8	20.9	16.2	21.3	92.0	11.3	2.4	26.8	100	97	13.6	-0.93	-0.76	18.2
CG12-360009	7.9	66.6	50.3	79.3	22.0	16.4	21.6	92.1	10.6	2.6	26.0	100	100	13.2	0.34	0.48	18.1
CP88-1165	9.2	76.7	46.6	77.7	22.3	16.8	22.2	91.9	10.2	2.5	26.1	85	98	12.5	1.72	1.48	17.9
CG12-606100	5.8	82.2	58.8	79.2	21.7	16.1	21.6	90.7	12.4	2.4	24.3	100	99	12.1	-0.62	-0.69	17.8
CG12-123	7.3	76.4	58.8	79.5	22.2	15.9	21.0	91.9	12.1	2.7	22.7	100	100	14.2	-1.11	-0.81	17.3
CG12-057019	7.6	81.1	59.0	79.3	23.0	16.5	21.9	91.2	11.5	2.8	28.5	100	97	11.2	1.52	1.57	17.0
CP72-2086	6.7	67.7	56.5	79.5	22.9	16.0	21.7	90.9	10.4	2.5	26.8	100	100	12.6	-0.26	-0.48	17.0
CG12-317044	7.1	65.9	55.3	78.8	22.5	16.6	21.8	92.4	11.0	2.6	25.9	100	95	12.2	0.96	1.02	17.0
CG12-124	6.0	80.3	58.8	79.2	23.5	16.5	21.9	91.7	13.0	2.5	24.1	100	100	13.4	0.97	0.73	17.0
CG12-332015	6.7	82.3	58.6	79.3	23.0	16.7	22.1	92.1	10.4	2.3	24.2	100	100	12.6	1.14	0.73	16.8
CG12-531020	7.1	83.8	53.0	78.8	21.2	16.7	22.1	91.9	12.0	2.5	23.2	100	91	12.0	1.21	1.20	16.7
CG12-318022	8.3	72.2	46.2	72.7	21.9	16.6	22.0	92.2	10.5	2.6	27.2	100	95	13.3	1.32	1.28	16.3
CG12-337009	7.3	67.3	51.6	77.8	22.3	16.4	21.5	92.5	11.1	2.7	26.1	100	94	12.8	0.39	0.61	16.3
Continuación Tabla 4.36																	
CG98-78	9.8	75.7	38.5	79.1	22.2	16.3	21.5	92.3	10.0	2.7	28.1	100	93	13.1	0.38	0.60	16.1
CG12-534006	7.4	83.8	58.4	79.3	21.6	16.0	21.3	91.2	12.3	2.3	25.1	100	100	10.8	-1.26	-1.35	16.0
CG12-626007	6.3	74.5	47.6	76.9	21.6	15.4	21.0	89.0	13.9	2.3	23.9	85	71	13.1	-2.83	-3.10	16.0
CG12-317020	8.5	67.1	44.3	76.6	21.2	15.9	21.6	90.1	12.7	2.4	27.1	100	99	10.7	-0.78	-0.92	15.7
CG12-317043	6.6	66.6	34.5	79.2	21.5	15.8	21.2	90.4	10.6	2.3	26.6	100	100	12.1	-1.93	-2.10	15.7
CG12-121	6.2	76.5	54.1	79.2	22.1	15.7	21.0	90.6	12.2	2.6	26.3	100	97	12.7	-1.61	-1.44	15.4
CG12-332013	7.4	84.4	58.8	79.4	22.9	16.7	21.9	92.8	10.6	2.3	24.0	100	100	12.7	0.82	0.50	14.5
CG12-570001	6.3	74.6	29.3	72.9	21.4	16.2	21.3	92.6	12.1	2.3	23.3	100	95	11.9	-1.03	-1.08	14.3
CG12-564013	7.2	74.6	58.4	77.8	21.5	15.9	21.3	90.4	11.3	2.4	27.9	85	98	11.8	-1.36	-1.36	13.8

ENTRY	peso5stalk_2	oqueleve_3	oqueinter_4	oqueleve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14	ESIM	RESIM	TAH
CG12-606012	5.5	75.2	46.2	79.4	23.3	17.0	22.2	92.9	11.4	2.3	23.2	100	95	11.9	1.86	1.53	13.0
CG12-637001	7.0	74.8	52.9	78.6	21.4	15.6	21.1	91.1	11.0	2.6	28.0	100	95	11.9	-1.70	-1.43	12.9

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueleve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

Tabla 4.37. Resumen de las Salidas de RINDSEL para los índices ESIM y RESIM para el conjunto de datos “medio_no_flor_1^a_soca”.

Índice	Estadístico	Valor del estadístico	Fracción superior (10% Variedades Seleccionadas)	Valor del índice
ESIM	Covarianza entre el índice y el VMG	0.190	CG12-512013	3.629
	Varianza del índice	0.492	CG12-333077	2.657
	Varianza del VMG	0.051	CG12-034098	2.469
	Correlación entre el índice y el VMG	0.999	CG12-330036	2.399
			CG12-606012	1.876
RESIM	Covarianza entre el índice y el VMG	0.191	CG12-512013	3.515
	Varianza del índice	0.431	CG12-333077	2.508
	Varianza del VMG	0.070	CG12-330036	2.350
	Correlación entre el índice y el VMG	0.999	CG12-034098	2.291
			CG12-057019	1.603

VMG: Valor de la Mejora Genética (Breeding value).

4.13.2.4 CPs como índices de desempeño para el conjunto de datos medio_no_flor_1^a_soca

Las funciones lineales generadas después de un ACP se emplearon aquí como índices de desempeño. En el acápite 4.10.2 se abordó con más detalle los principales resultados del análisis. El total de la variabilidad explicada por las tres primeras CPs para este conjunto de datos fue de ~46% (Tabla 4.12). La CP1 (19.8%) estuvo asociada a las variables de concentración de azúcar, la CP2 (14.9%) se asoció a las variables **pith_13**, **pop_9**, **flower_12**, **peso5stalk_2**, **diam_11** y **oqueleve_3**. Es decir, la CP2 se asoció a una combinación de variables relacionadas con la acumulación de biomasa y calidad interna del tallo. Por otro lado, la CP3 (11.2%) se asoció prácticamente solo con variables relacionadas con la

acumulación de biomasa (**alt_10**, **diam_11**, **peso5stalk_2**, y **pop_9**). Se siguió la propuesta de selección anidada según los *scores* de las CPs, análogamente a cómo se hizo para el conjunto de datos anteriores.

La CP2 presentó la heredabilidad más alta (Tabla 4.38), seguida por la CP3. La CP1 presentó el valor más bajo de heredabilidad, lo cual concuerda con los valores individuales de heredabilidad de las variables relacionadas con la concentración de azúcar asociados con el ensayo de variedades **No Flor** del estrato medio (Tabla 4.7). Sin embargo, esta CP1 fue la que más variabilidad acumuló y contiene a las variables que, en orden de prioridad, debieran ser consideradas prioritarias para realizar el ejercicio de selección.

Tabla 4.38. Resumen de indicadores para los índices CP1, CP2 y CP3 para el conjunto de datos medio_no_flor_1ª_soca.

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY)	Varianza ambiental (RESIDUOS)	H ²	Valor estimado de Mejora genética =H ² *(ENTRY + RESIDUOS)
CP1	0	0.92145	1.73532	0.515	2.073
CP2	0	1.41519	0.66538	0.810	2.088
CP3	0	0.77193	0.57249	0.729	1.005

H²= heredabilidad en el sentido amplio

En la Tabla 4.39 se presentan los valores (*scores*) para la fracción superior del 25% siempre, en principio, con respecto a la CP1, seguidos de los mejores *scores* para la CP2 y, seguidamente, lo concerniente a la CP3. Se observa en dicha tabla que los genotipos priorizados en cada CP fueron diferentes, excepto para el genotipo CG12-MSING#4, el cual figuró como un buen genotipo en las tres CPs.

Tabla 4.39. Clasificación de los genotipos (fracción del -25%) según los *scores* de CP1, CP2 y CP3.

Genotipos según la CP1	<i>scores</i>	Genotipos según la CP2	<i>scores</i>	Genotipos según la CP3	<i>scores</i>
CG12-606012	-2.443	CG12-126	-2.367	CG12-229005	2.330
CG12-512013	-2.296	CG12-MSING#4	-2.226	CG12-MSING#4	2.064
CG98-78	-1.889	CG12-337009	-2.141	CG12-318012	1.903
CG12-324187	-1.880	CG12-132	-1.385	CG12-606102	1.566
CG12-627037	-1.413	CG12-324156	-1.360	CG12-324147	1.505
CG12-332015	-1.407	CG12-317020	-0.941	CG12-317044	0.970
CG12-318012	-1.399	CG12-229005	-0.936	CG12-627037	0.861
CG12-333077	-1.395	CG12-318012	-0.897	CG12-135	0.818
CG12-330036	-1.282	CG98-78	-0.895	CG12-629011	0.783
CG12-008054	-1.270	CG12-317044	-0.892	CG12-324156	0.780
CG12-123	-1.226	CG12-317043	-0.851	CG12-132	0.763
CG12-MSING#4	-1.144	CG12-318022	-0.822	CG12-333077	0.760
CG12-360009	-1.078	CG12-512009	-0.784	CG12-057019	0.759
CG12-324147	-1.058	CG12-130	-0.778	CG12-130	0.715

El Análisis de Varianza (ANOVA) mostró que las diferencias de las medias con respecto a la CP1 son una función de la variación genotípica (variedades) ($Pr > F = 0.0175$) (Tabla 4.17). Con este valor se procedió a una comparación de promedios por pares ajustada por Tukey, la cual se presentó en la Figura 4.24, donde se mostraron los resultados relacionados con los Análisis de Varianza de las CPs.

En la Figura 4.25 se mostró que las diferencias entre genotipos fueron solamente entre el genotipo CG12-130 y los últimos tres genotipos, los cuales presentaron los valores (*scores*) más bajos para la CP1. En este sentido, en la fracción superior del 25% de este conjunto de datos, los genotipos no presentaron diferencias entre sí, por lo que simplemente se seleccionaron los primeros mejores 14 genotipos (25%) (Tabla 4.39), sobre los cuales se realizó el ejercicio de selección final con respecto a las CPs.

Para ello, se realizó un ejercicio de selección anidada solamente con base en las CP1 y CP3, debido a que las cargas en la CP2 no estuvieron bien definidas con alguno de los grupos de variables (azúcar, biomasa, y calidad interna del tallo).

Una vez más, se nota que cada CP ordena las variedades según la importancia de las variables asociadas según los *scores* para la fracción superior del 25% (*eigenvectors*) (Tabla 4.40). El orden de importancia para realizar una selección anidada viene dado por la varianza

acumulada de cada CP. En primer lugar se encontró la CP1 por la importancia que tiene la concentración de azúcar (Figura 4.20 y 4.21). Le siguió la CP2, la cual no presentó un orden claro sobre las variables que estuvieron más asociadas, **pith_13**, **pop_9**, **flower_12**, **peso5stalk_2**, **oqueleve_3** y **diam_11**. En todo caso es una combinación de variables relacionadas con la biomasa y la calidad interna del tallo. Por tanto, en este conjunto de datos se realizó la selección anidada de la siguiente manera: selección en la CP1 y, dentro de estas selecciones, se eligieron los genotipos con los mejores *scores* correspondientes en la CP2; y por último, dentro de estas selecciones, se seleccionaron las variedades con los mejores *scores* en la CP3.

Tanto la CP2 como la CP3 incluyeron cargas importantes de las variables relacionadas con la acumulación de biomasa, por lo que tuvieron un peso importante en las selecciones. Las cargas en las CP1 y CP2 fueron principalmente de signo negativo (Tabla 4.13), por lo que, en estas componentes, los mejores genotipos fueron aquellos que presentaron los *scores* más bajos. Las selecciones en la CP1 se realizaron siguiendo, inicialmente, un análisis de varianza para un modelo lineal mixto, donde el término del mismo para los genotipos (ENTRY) se consideró como fijo, y el resto de los términos del modelo fueron considerados como términos aleatorios. Se comprobó que los datos siguen una tendencia apropiada para su tratamiento en un Análisis de Varianza (Figura 4.34).

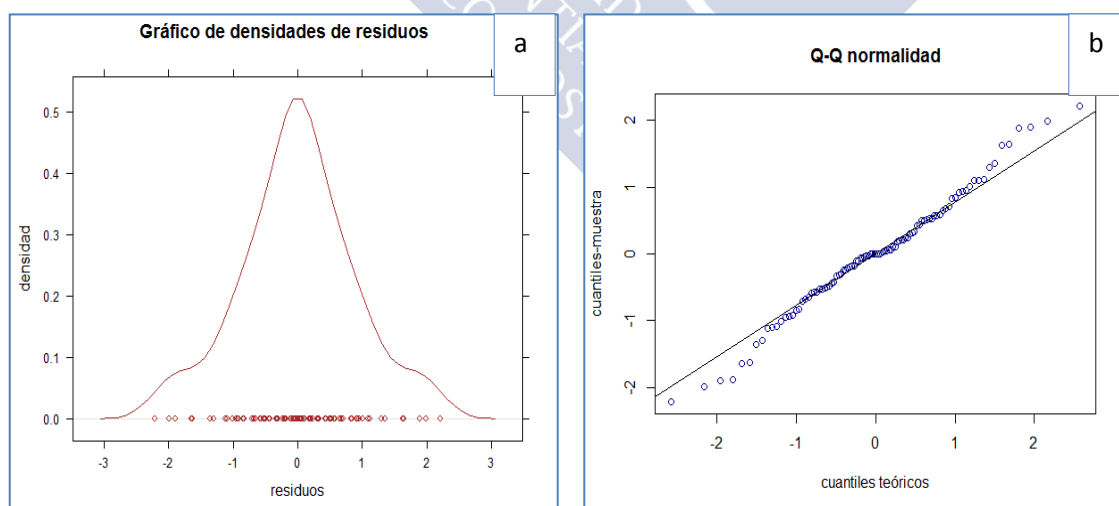


Figura 4.34. Propiedades sobre normalidad de la serie de datos *medio_no_flor_1ª_soca*.

Tabla 4.40. Fracción superior de ~25% inicialmente seleccionada por la CP1 para el conjunto de datos medio_no_flor_1ª_soca. Separación de medias mediante el método de Tukey para la CP1. En verde, genotipos resultantes de la selección anidada a partir de las CP1, CP2 y CP3.

Genotipos (ENTRY)	Tukey CP1	CP1	Genotipos (ENTRY)	CP2	Genotipos (ENTRY)	CP3
CG12-606012	B	-2.443	CG12-MSING#4	-2.226	CG12-MSING#4	2.064
CG12-512013	B	-2.296	CG12-318012	-0.897	CG12-318012	1.903
CG98-78	AB	-1.889	CG98-78	-0.895	CG12-324147	1.505
CG12-324187	AB	-1.880	CG12-008054	-0.595	CG12-627037	0.861
CG12-627037	AB	-1.413	CG12-324187	-0.224	CG12-333077	0.760
CG12-332015	AB	-1.407	CG12-330036	-0.016		
CG12-318012	AB	-1.399	CG12-360009	0.035		
CG12-333077	AB	-1.395	CG12-606012	0.164		
CG12-330036	AB	-1.282				
CG12-008054	AB	-1.270				
CG12-123	AB	-1.226				
CG12-MSING#4	AB	-1.144				
CG12-360009	AB	-1.078				
CG12-324147	AB	-1.058				

4.13.2.5 TAH (t De Azúcar Por ha) y BLUP (*Best Linear Unbiased Prediction*)

El valor de TAH tuvo un valor intermedio de heredabilidad (Tabla 4.41), con un valor genotípico del 34.2%. El resto del valor puede ser imputable al ambiente. En todo caso, TAH se presentó como una buena opción de la valoración genética de los genotipos en evaluación, según el VMG.

Tabla 4.41. Resumen de indicadores de los índices TAH para el conjunto de datos medio_no_flor_1ª_soca.

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY)	Varianza ambiental (RESIDUOS)	H ²	Valor de Mejora genética (VMG)=H ² *(ENTR Y + RESIDUOS)
TAH	18.66 ± 3.33	3.6034	13.8789	0.342	5.97

H²= heredabilidad en el sentido amplio

La varianza ambiental presentó un valor muy superior al de la varianza genotípica para el TAH (Tabla 4.41), lo que redundó en un valor bajo de H² y de VMG. Esto, en general, se reflejó en una alta variabilidad de respuesta a través de todo el conjunto de genotipos evaluados, como producto de una alta incidencia del ambiente.

Si bien el valor fenotípico medio de TAH para este conjunto de datos **medio_no_flor_1ª_soca** fue similar al del anterior conjunto de datos, su desviación fue menor, lo cual significa que, en general, presentó valores de TAH más estables y presentó un mayor número de genotipos con valores aceptables, comparado con el conjunto de datos **medio_flor_plantía**. En este sentido, es posible aceptar que el desempeño fenotípico del experimento **medio_no_flor_1ª_soca** fue superior al que se presentó anteriormente (**medio_flor_plantía**), ya que la frecuencia absoluta de genotipos ubicados en el rango de 16.57 TAH a 31.2 TAH fue de 36, en tanto que para el conjunto de datos **medio_flor_plantía**) en el rango de 15.02 TAH a 25.9 TAH fue de 22 (Tabla 4.42). La premisa fue que, si el experimento resultó en más genotipos con mejor desempeño, el sitio y la época de evaluación, combinados, presentaron mejores condiciones de evaluación y selección

Tabla 4.42. Frecuencias de genotipos evaluados en los tres conjuntos de datos según sus valores de TAH.

Conjunto de datos	Clase	Límite inferior	Límite superior	Valor medio	Frecuencia absoluta	Frecuencia Relativa
medio_flor_plantía	1	11.40	15.02	13.21	8	0.27
	2	15.02	18.65	16.83	8	0.27
	3	18.65	22.27	20.46	6	0.20
	4	22.27	25.89	24.08	8	0.27
medio_no_flor_1a_soca	1	12.91	16.57	14.74	13	0.27
	2	16.57	20.24	18.41	21	0.43
	3	20.24	23.91	22.08	13	0.27
	4	23.91	27.58	25.74	1	0.02
	5	27.58	31.25	29.41	1	0.02
litoral_no_flor_plantía	1	11.72	16.73	14.22	5	0.12
	2	16.73	21.74	19.24	14	0.34
	3	21.74	26.75	24.25	11	0.27
	4	26.75	31.77	29.26	8	0.20
	5	31.77	36.78	34.27	3	0.07

En un Análisis de Varianza para la variable TAH, para un modelo lineal mixto, no se detectaron diferencias estadísticas en los valores de la variable de respuesta ($F=1.562$; $Pr>F=0.063$, NS), por lo que no se realizó la separación de promedios correspondiente. Por lo tanto, se testó la selección por medio de BLUP, puesto que podría ser más apropiada en este caso en particular.

Los mejores estimadores lineales insesgados (BLUPs) respecto a la variable TAH calculados para cada genotipo mostraron una distribución de 27 genotipos con valores inferiores a cero, y 22 genotipos con valores superiores a cero (Figura 4.33.a). Es decir, que la mayoría de los genotipos en este conjunto de datos estuvieron asociados a valores bajos de TAH. Un genotipo presentó un valor extraordinario de 31.24 TAH, lo que se vio reflejado con un valor más alto de BLUP (Figura 4.35.a y 4.35.b) y, además, con un rango de frecuencias (Figura 4.35.a) muy diferente al resto de materiales. Los valores del TAH fuera de la línea diagonal de asociación (la línea diagonal correspondió a la media de la variable original) con los BLUPs fueron dos, sin embargo, éstos fueron corregidos por el BLUP correspondiente. El ajuste por medio de BLUP se dio cuando el estimador (BLUP) respectivo de un valor alejado de la media acercó un poco más el nuevo valor (BLUP) a la media, respecto al valor original de la variable TAH; de esta forma se permite que este nuevo valor sea un estimador del valor genético del material vegetal en cuestión; y, a su vez, permite que su valoración sea más acorde a la tendencia general.

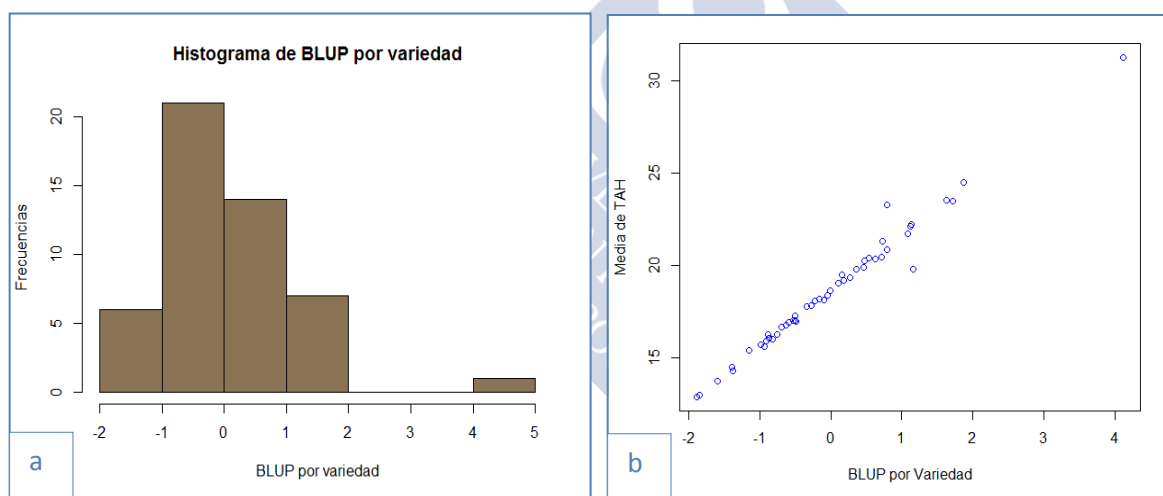


Figura 4.35. a) Distribución de frecuencias absolutas de los BLUPs. b) BLUPs ajustados a la media de TAH.

En la Tabla 4.43 se presenta un listado descendente del 25% de la fracción superior en este conjunto de datos en función de los valores de BLUP para el conjunto de datos **medio_no_flor_1ª_soca**. El ordenamiento de BLUP presentó una selección del 10% de la fracción superior diferente de las que hubiesen sido planteadas por el TAH. El genotipo CG12-533002 (22.25 TAH), por ejemplo, hubiese sido seleccionado en el esquema de TAH; sin embargo, después de la corrección realizada por BLUP, la sugerencia de selección fue el

genotipo CG12-034098 (19.81 TAH). En este conjunto de datos se observó un ejemplo más del efecto de BLUP y las sugerencias de selección diferentes a TAH, de manera similar a lo que ocurrió en el conjunto de datos previamente presentado. Un tema adicional en este conjunto de datos fue que se encontró que el mejor testigo fue la variedad CP88-1165 (17.86 TAH), la cual presentó un valor bajo de BLUP y se encontró ubicada en el segundo cuartil de la población en estudio (en el rango -0.384 a 1.118), mientras que todas las selecciones se ubicaron en el último cuartil (2.621 a 4.123) (Tabla 4.44). Esto indica que fue posible identificar una alta frecuencia genotipos que superaron en desempeño a los testigos propuestos en este conjunto de datos.

Tabla 4.43. Ordenamiento descendente de la fracción superior (~25%) de genotipos según los valores de BLUP para medio_no_flor_1ª_soca.

Genotipo	BLUP	TAH
CG12-229005	4.12	31.25
CG12-130	1.87	24.50
CG12-MSING#4	1.72	23.47
CG12-132	1.63	23.53
CG12-034098	1.16	19.81
CG12-533002	1.14	22.25
CG12-318012	1.13	22.11
CG12-008054	1.08	21.71
CG12-627037	0.80	23.27
CG12-324147	0.79	20.84
CG12-135	0.72	21.30
CG12-122	0.72	20.48
CG12-126	0.62	20.36
CP88-1165 (Testigo)	-0.27	17.86

Tabla 4.44. Frecuencias de distribución de BLUP por cuartiles.

Cuartil	Límite inferior	Límite superior	Mediana	Frecuencia absoluta	Frecuencia relativa
1	-1.887	-0.384	-1.135	20	0.408
2	-0.384	1.118	0.367	22	0.449
3	1.118	2.621	1.870	6	0.122
4	2.621	4.123	3.372	1	0.020

4.13.2.6 Regresión logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos medio_no_flor_1ª_soca

Al considerar la regresión Logit para el modelo correspondiente (ver sección de Materiales y Métodos), se observó en la mayoría de los casos que los índices presentaron valores bajos de asociación con la base de comparación, tanto en los valores de las funciones logísticas (Tabla 4.45) como en los gráficos de asociación entre seis de los ocho índices evaluados (Figura 4.36), excepto para los índices TAH y BLUP, los cuales presentaron modelos con alta significación estadística y valores intermedios de Pseudo r^2 ; así mismo, presentaron un mejor ajuste de la curva característica de una regresión Logit (Figura 4.34).

Tabla 4.45. Valores de Pseudo r^2 de McFadden de una regresión Logit para el modelo general $\ln[p/(1-p)] = a + bX + e$ para el conjunto de datos medio_no_flor_1ª_soca.

índice	Modelo	Prueba de Significación $Pr(>x^2)$	Pseudo- r^2 con la base selección por competencia	Eficiencia con respecto al r^2 más alto
TAH	$\ln[Si/(1-Si)] = -11.9895 + 0.7167 \text{ TAH}$	<0.0001	0.341	1.00
BLUP	$\ln[Si/(1-Si)] = 1.3679 + 2.1268 \text{ BLUP}$	<0.0001	0.337	0.99
CP2	$\ln[Si/(1-Si)] = 0.7360 - 0.2270 \text{ CP2}$	0.321	0.016	0.20
ESIM	$\ln[Si/(1-Si)] = 0.7897 + 0.4295 \text{ ESIM}$	0.054	0.060	0.20
RESIM	$\ln[Si/(1-Si)] = 0.7986 + 0.4796 \text{ RESIM}$	0.039	0.069	0.18
CP3	$\ln[Si/(1-Si)] = 0.7518 + 0.3954 \text{ CP3}$	0.196	0.027	0.08
CP1	$\ln[Si/(1-Si)] = 0.7875 - 0.4823 \text{ CP1}$	0.042	0.067	0.05
KN	$\ln[Si/(1-Si)] = 0.7296 + 0.4688 \text{ KN}$	0.550	0.006	0.02
Smith	$\ln[Si/(1-Si)] = 0.7244 + 0.1890 \text{ Smith}$	0.870	0.000	0.00

a=intercepto, b=coeficiente, X=Índice, e=término del error, p=probabilidad de ocurrencia de "sí", 1-p=probabilidad de ocurrencia de "no".

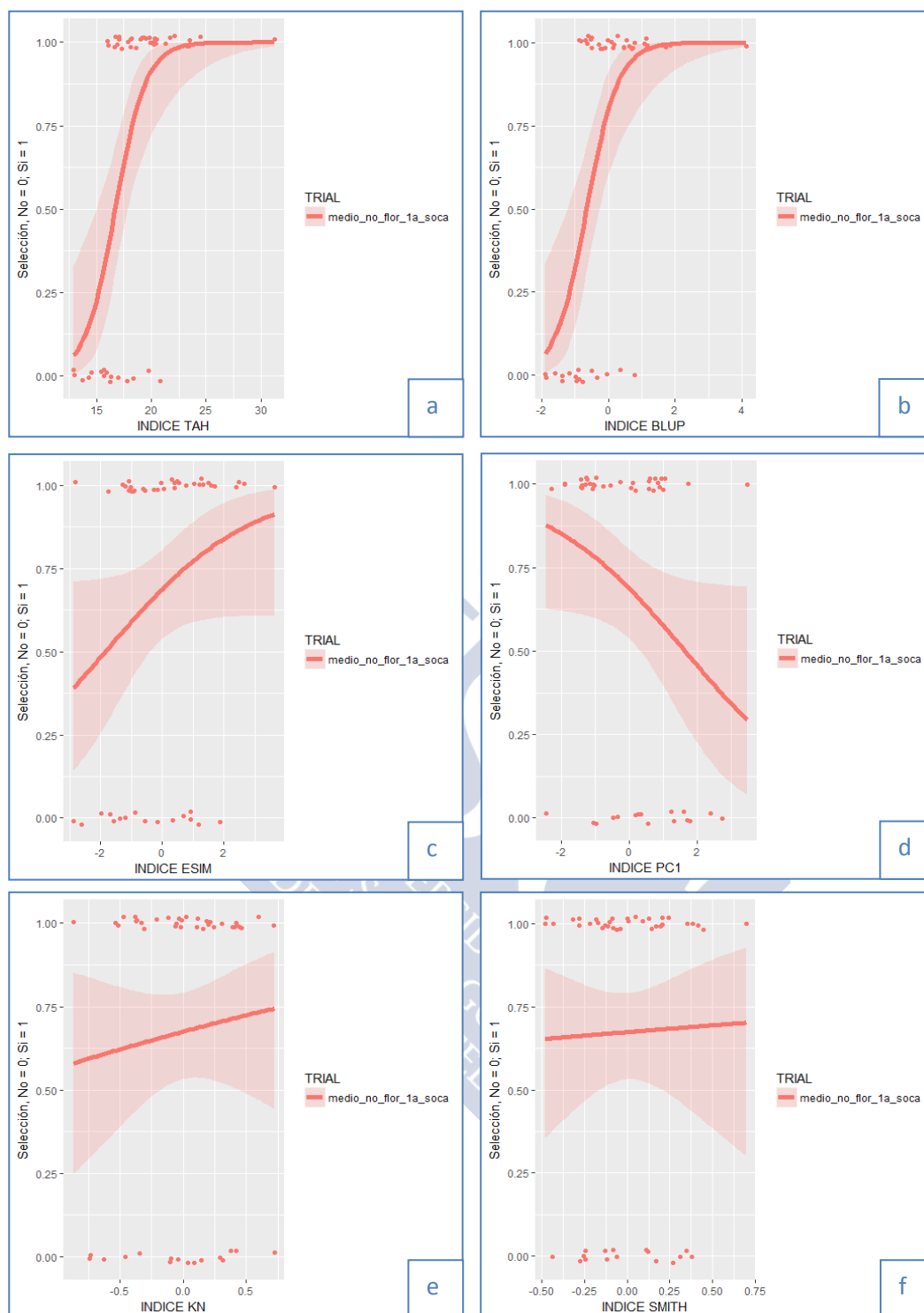


Figura 4.36. Gráficos de las funciones logit en el conjunto de datos medio_no_flor_1^a_soca para los índices: a) TAH , b) BLUP, c) ESIM, d) CP1, e) KN y f) Smith.

4.13.3 Índices de selección para el conjunto de datos litoral no flor plantía (litoral_no_flor_plantía)

4.13.3.1 Análisis de correlación entre índices

El conjunto de datos **litoral_no_flor_plantía** fue el que presentó los promedios de desempeño fenotípico más elevados, especialmente en los aspectos de acumulación de biomasa. Este aspecto se reflejó en la producción final de azúcar medida en TAH (Tabla 4.2).

En este conjunto de datos, el valor genético (VMG) se vio influido por variables de calidad del tallo como **oqueleve_3**, **oqueinter_4**, **oqueseve_5**, **flower_12**, **diam_11** y **pop_9**, las cuales presentaron valores de heredabilidad relativamente altos (Tabla 4.4). Esta relación hizo que el valor genético de los materiales en evaluación se extendiese hasta esas variables. El índice señaló correctamente estas variables, pero su ponderación dentro del índice fue muy alta, en detrimento de otras variables como el Peso de cinco tallos, y las variables de concentración de sacarosa. En este sentido, el índice no reflejó las necesidades del fitomejorador. En todo caso, pudiera ser aplicable si se considerase un menor número de variables y que estuvieran más o menos correlacionadas entre sí, como las componentes del rendimiento; sin embargo, lo anterior no sería una solución aceptable por cuanto las 14 variables en estudio son importantes.

Los resultados obtenidos indicaron que es posible asegurar que los índices TAH y BLUP, así como **ESIM** y **RESIM**, fueron intercambiables entre cada par (Tabla 4.46). En este sentido, el uso de BLUP puede ser más apropiado en virtud de que se considera un estimador insesgado, y dota a los genotipos en evaluación de un valor diferencial robusto que puede ser empleado para la discriminación entre ellos y, en consecuencia, contribuir con la selección de genotipos. Así mismo, la correspondencia entre **ESIM** y **RESIM** puede ser motivo de uso intercambiable. Las circunstancias pueden obligar al uso de uno u otro, y puede estar derivado de las restricciones que se impongan en **RESIM**. Esta asociación entre estos dos índices está relacionada, probablemente, con las variables que imponen la restricción, tal como ya se indicó en los conjuntos de datos anteriores. Seguramente, las variables que impusieron restricciones en **RESIM** fueron también de un alto peso en la función que se determinó para **ESIM**.

Las correlaciones entre **ESIM** y **RESIM** con las CP1 y CP2 pueden estar asociadas a las variables que gobiernan estos índices, las cuales estuvieron relacionadas con la sacarosa para la CP1, y la combinación de variables relacionadas con la acumulación de biomasa y la calidad interna de los tallos (Figura 4.20). Smith, se correlacionó negativamente con los índices **ESIM** y **RESIM**, y en menor cuantía de forma negativa con la CP1 y CP3 (Tabla 4.46).

Tabla 4.46. Correlación de Pearson y Significación estadística entre los valores de los índices obtenidos para el conjunto de datos litoral_no_flor_plantía.

	TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3
TAH		****	***	***	NS	*	***	**	NS
BLUP	1.00		***	****	NS	*	***	**	NS
ESIM	0.55	0.56		****	NS	***	**	NS	NS
RESIM	0.57	0.58	0.99		NS	**	***	NS	NS
KN	0.21	0.21	0.19	0.11		****	NS	NS	*
Smith	-0.35	-0.36	-0.50	-0.43	-0.75		NS	*	*
CP1	-0.52	-0.50	-0.48	-0.53	0.26	-0.15		NS	NS
CP2	-0.47	-0.47	-0.26	-0.28	-0.12	0.33	0.06		NS
CP3	0.27	0.25	0.13	0.08	0.32	-0.33	0.08	0.19	

Códigos para la significación estadística: ****, $p \leq 0.0001$; ***, $p \leq 0.001$; **, $p \leq 0.01$; *, $p \leq 0.05$; NS, no significativo

4.13.3.2 Índices “Smith” y “Restrictive Kempthorne And Nordskog Selection Index” (Índice KN)

En la Tabla 4.47 se presentan los promedios para las variables estudiadas así como los *scores* de los índices Smith (en orden descendente), KN y la referencia de los valores de TAH. En general, se presentaron valores elevados para todas las variables según se apuntó anteriormente, particularmente para la variable/índice TAH. Smith y KN fueron índices derivados de trece variables, ya que **flower_12** no fue incluida para la estimación de los índices debido a que su variación fue prácticamente cero, como puede observarse en la Tabla 4.47 en su valor transformado ($100 = 0$). Si bien la correlación de estos índices fue baja (-0.18 , NS), los valores de ambos índices se fundamentaron en prácticamente las mismas variables (**brixprom_1**, **peso5stalk_2**, **pol_6**, **pith_13**, **fiber_14**). Las propuestas de selección del 10% de la fracción superior, para ambos índices, fueron diferentes.

Tabla 4.47. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo litoral_no_flor_plantía.

Genotipos (ENTRY)	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14	Smith	KN	TAH
CG12-317021	7.5	85.6	58.8	79.3	19.2	16.2	22.1	90.1	13.0	2.1	26.1	100	46	12.3	2.31	1.06	20.6
CG12-036037	6.6	77.2	56.0	78.9	23.3	16.4	21.6	92.4	13.9	2.2	25.5	100	100	12.1	1.91	0.59	16.8
CG12-113	6.5	84.9	59.1	79.4	22.7	16.5	21.8	91.9	15.1	2.3	25.4	100	100	12.5	1.70	0.65	24.1
CG12-107	5.9	81.5	58.8	79.3	21.7	15.3	20.9	89.2	14.5	2.2	22.9	100	100	12.1	1.67	0.85	17.9
CG12-317012	7.5	71.1	47.1	79.6	21.0	15.5	20.7	91.0	15.8	2.4	25.1	100	100	13.2	0.85	0.30	20.4
CG12-111	7.8	86.0	59.4	79.5	21.2	15.9	20.7	89.9	14.4	2.4	26.1	100	99	12.2	0.68	-0.17	36.8
CG12-104	6.3	77.1	58.5	78.3	21.6	15.2	20.1	91.1	14.9	2.4	23.4	100	97	11.9	0.68	0.12	25.1
CG12-532007	8.8	72.3	58.4	79.6	21.5	16.4	21.7	92.3	13.1	2.4	27.7	100	100	12.1	0.68	0.00	27.2
CG12-106	4.8	86.5	58.8	79.3	21.0	14.6	19.7	89.9	12.5	2.4	19.5	100	99	14.0	0.61	0.40	11.7
CG12-101	7.1	77.6	58.8	79.4	22.4	15.4	21.0	90.8	15.5	2.4	24.3	100	85	13.0	0.56	0.61	23.4
CG12-108	7.1	86.5	58.8	79.3	20.4	15.0	20.1	90.0	16.0	2.4	23.9	100	100	13.7	0.54	0.38	22.4
CG12-060016	5.3	69.8	52.0	79.3	20.1	15.1	19.5	92.2	15.0	2.4	22.9	100	73	12.1	0.50	-0.29	15.7
CP88-1165	7.1	58.1	58.8	79.3	22.0	15.5	20.1	93.8	14.5	2.4	26.7	90	100	10.6	0.46	-0.05	21.2
CG12-018003	5.7	77.1	59.4	78.6	22.2	15.0	20.1	90.6	14.9	2.4	22.2	100	97	11.3	0.38	0.09	14.3
CG12-339052	6.4	74.0	58.8	77.7	20.6	14.7	20.3	89.0	17.0	2.4	22.0	100	71	12.2	0.33	0.54	21.3
CG12-311020	8.5	67.7	57.8	78.9	21.0	15.2	20.8	88.7	13.0	2.4	26.8	100	100	12.5	0.25	0.20	24.0
CG12-318005	6.4	82.9	58.5	79.5	20.4	13.9	19.3	88.5	16.6	2.4	21.1	100	86	11.7	0.21	0.37	21.2
CG12-606129	8.4	85.6	55.7	74.1	21.1	15.2	20.4	89.2	13.5	2.5	26.1	100	78	12.9	0.18	0.08	20.0
CP72-2086	8.9	80.6	57.8	79.1	21.8	16.9	22.2	92.8	16.4	2.5	26.9	95	100	12.0	0.11	-0.02	28.6
CG12-620015	7.1	42.7	44.0	78.9	21.1	14.6	19.7	89.5	15.3	2.4	25.7	100	100	11.7	0.09	-0.06	21.8
CG12-102	7.3	85.6	57.7	78.4	20.6	14.7	20.2	88.6	14.3	2.4	25.3	100	78	12.3	0.07	0.33	19.1
CG12-330012	6.6	74.1	57.5	77.6	22.7	16.4	21.7	91.4	12.0	2.5	25.3	100	44	12.4	0.02	-0.07	17.5
CG12-030014	8.7	75.7	57.4	76.9	22.8	16.1	21.5	91.2	12.4	2.5	23.5	100	100	13.5	0.01	-0.11	21.1
CG12-112	6.6	73.1	56.5	79.6	22.1	15.3	20.6	89.9	14.0	2.5	23.3	100	82	11.6	-0.05	-0.02	21.6
CG12-106001	9.0	80.9	58.0	79.3	22.7	16.1	21.0	92.9	15.0	2.6	27.9	100	100	12.7	-0.26	-0.29	29.0
CG12-005031	8.6	81.0	58.2	79.2	21.0	15.5	20.7	91.0	14.9	2.6	25.4	100	100	11.1	-0.33	-0.30	20.8
CG12-110	6.8	83.5	56.9	79.2	20.7	15.5	20.7	91.5	16.9	2.6	22.9	100	99	13.7	-0.35	0.03	22.0
CG12-526031	9.7	75.5	51.6	79.3	22.8	16.9	22.1	92.2	12.7	2.6	26.6	100	96	11.3	-0.41	-0.73	22.8
CG12-276003	8.3	83.8	56.2	79.1	20.7	15.4	20.5	88.5	12.1	2.5	27.1	100	93	11.9	-0.41	-0.46	14.7
CG12-103	7.5	79.5	46.7	79.3	21.8	15.0	20.2	89.9	9.5	2.6	23.0	100	82	15.0	-0.49	-0.13	14.3
CG12-116	8.9	70.3	54.6	79.3	22.0	17.1	22.5	92.4	16.5	2.7	24.7	100	88	11.4	-0.51	-0.47	33.8
CG12-105	7.0	83.2	50.8	65.6	21.3	14.5	20.0	87.6	16.4	2.6	24.0	100	99	11.3	-0.83	-0.08	22.7
SP79-1287	9.6	75.1	58.4	79.4	20.5	15.5	20.9	90.3	14.8	2.6	27.7	95	100	11.7	-0.85	-0.20	29.2
CG12-114	6.8	86.5	58.8	81.0	22.5	15.3	20.8	89.6	16.7	2.6	22.0	100	100.0	12.1	-0.96	-0.08	22.0
CG12-326030	7.8	79.9	59.1	79.1	21.2	15.5	20.9	90.3	15.6	2.7	23.0	100	82.3	12.9	-1.05	-0.25	28.3
CG12-339025	8.6	80.3	58.1	79.5	22.2	14.6	19.9	90.5	14.8	2.7	22.4	100	100.0	11.9	-1.20	-0.38	19.0
CG12-311019	9.5	76.9	52.5	73.9	22.5	15.2	21.0	89.2	13.9	2.7	27.0	100	99.3	11.4	-1.24	-0.30	29.4
CG12-115	8.3	77.1	58.8	79.3	19.6	13.0	17.7	89.4	18.5	2.7	28.3	100	100.0	12.6	-1.24	-0.19	25.8
CG12-324094	8.5	52.3	49.4	79.1	21.4	15.3	20.3	90.1	15.8	2.7	26.3	100	95.5	10.3	-1.29	-0.91	32.7
CG12-109	7.8	81.3	55.7	79.5	20.3	15.2	20.7	89.3	15.6	2.7	24.8	100	93.3	13.3	-1.66	-0.35	27.3
SP71-6161	8.3	86.5	58.8	79.3	21.3	16.3	21.7	91.2	16.0	2.8	23.0	100	100.0	11.4	-1.68	-0.71	29.2

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

El índice Smith presentó un valor de correlación de 1.00 con VGM (Tabla 4.48) y de 0.67 con KN. Estas diferencias en la valoración genética de los materiales pudieron haber dado lugar a selecciones diferentes para cada índice. Las selecciones no correspondieron con el mejor desempeño de los genotipos en términos de acumulación de biomasa y concentración de azúcar, es decir, TAH (Tabla 4.45). En este conjunto de datos, el valor genético (VMG) se vio influido por variables de calidad del tallo **oqueleve_3**, **oqueinter_4**, **oqueseve_5**, **flower_12**, **pith_13**; y otras variables como **pop_9** y **diam_11**, las cuales presentaron valores entre intermedios y elevados de heredabilidad para este conjunto de datos (Tabla 4.7). Esta relación hizo que el valor genético de los materiales en evaluación se extendiese hasta esas variables. El índice señaló correctamente estas variables, pero su ponderación dentro del índice fue muy alta, en detrimento de otras variables como el **peso5stalk_2** y las variables de concentración de sacarosa (**pol_6** y **brixprom_1**). En este sentido, el índice no reflejó las necesidades de fitomejoramiento. En todo caso, pudiera ser aplicable si se considerase un menor número de variables, las cuales pudieran estar más correlacionadas entre sí; sin embargo, lo anterior no sería una solución aceptable porque las 14 variables en estudio son importantes.

Una situación similar se encontró para el índice KN, si bien la correlación con el VMG fue menor debido a la restricción a las variables de importancia según lo que se indicó antes; y, además, porque no incluyen a las variables que se indicaron en el párrafo anterior. Por lo tanto, la selección se ajustó un poco mejor a las necesidades del fitomejorador. En este sentido, pareció señalar selecciones de mejor calidad.

Con respecto a Smith y KN puede decirse que los índices no parecieron adaptarse adecuadamente a las necesidades de la mejora genética en caña de azúcar, ya que los desempeños de las selecciones, en términos de producción de azúcar, no fueron los mejores según lo que se puede observar en la Tabla 4.47. Uno de los principales argumentos de selección tradicionales en caña de azúcar es que los genotipos superiores presenten altos desempeños en productividad de azúcar. Luego se evalúan el resto de elementos, y se decide la selección final.

Tabla 4.48. Resumen de las Salidas de RINDSEL para los índices Smith y KN para el conjunto de datos litoral_no_flor_plantía.

Índice	Estadístico	Valor del estadístico	Fracción superior (10% Variedades Seleccionadas)	Valor del índice
Smith	Covarianza entre el índice y el VMG	6.782	CG12-113	1.588
	Varianza del índice	0.080	CG12-111	0.736
	Varianza del VMG	299.593	CG12-532007	0.540
	Correlación entre el índice y el VMG	0.999	CG12-104	0.628
			CP88-1165 (testigo)	0.644
			SP71-6161(testigo)	-1.618
			SP79-1287(testigo)	-0.721
KN	Covarianza entre el índice y el VMG	4.418	CG12-101	0.252
	Varianza del índice	0.147	CG12-311019	0.513
	Varianza del VMG	299.593	SP71-6161	0.454
	Correlación entre el índice y el VMG	0.665	CG12-109	0.571
			CP88-1165(testigo)	-0.406
			SP71-6161(testigo)	0.454
			SP79-1287(testigo)	0.097

VMG: Valor de la Mejora Genética (*Breeding value*).

4.13.3.3 *Eigenanalysis-based Selection Index Method (ESIM)* y *Restrictive Eigenanalysis-based Selection Index Method (RESIM)*

La Tabla 4.49 presenta los valores promedio de las variables de estudio, más los *scores* de **ESIM** y **RESIM** y la referencia del TAH como medida generalizada de desempeño. **ESIM** y **RESIM** son índices derivados de trece variables, ya que **flower_12** no fue incluida para la estimación de los índices debido a que su variación fue prácticamente cero como puede observarse en la Tabla 4.47 en su valor transformado ($100 = 0$).

ESIM y **RESIM** presentaron valores altos de correlación con el VGM (Tabla 4.50). Las selecciones propuestas por ambos índices fueron prácticamente las mismas, y solo difirieron con respecto a las variedades CG12-109 para **ESIM** y CG12-311019 para **RESIM**. Los valores para ambos índices estuvieron asociados a niveles medios y elevados en las variables relacionadas con la biomasa, concentración de azúcar, y calidad del tallo, por lo que las selecciones se presentan como apropiadas; el nivel de correlación con la base de comparación es bajo.

Tabla 4.49. Valores promedio de 14 variables de estudio en cada genotipo en evaluación y los valores correspondientes de los índices Smith, KN y TAH para el ensayo litoral_no_flor_plantía.

Genotipos (ENTRY)	peso5stalk_2	oqueleve_3	oqueinter_4	oqueseve_5	brixprom_1	pol_6	brixjugo_7	purity_8	pop_9	alt_10	diam_11	flower_12	pith_13	fiber_14	ESM	RESM	TAH
CG12-317021	7.5	85.6	58.8	79.3	19.2	16.2	22.1	90.1	13.0	2.1	26.1	100	46	12.3	2.23	2.60	20.6
CG12-036037	6.6	77.2	56.0	78.9	23.3	16.4	21.6	92.4	13.9	2.2	25.5	100	100	12.1	0.01	0.12	16.8
CG12-113	6.5	84.9	59.1	79.4	22.7	16.5	21.8	91.9	15.1	2.3	25.4	100	100	12.5	1.95	2.30	24.1
CG12-107	5.9	81.5	58.8	79.3	21.7	15.3	20.9	89.2	14.5	2.2	22.9	100	100	12.1	-0.21	-0.09	17.9
CG12-317012	7.5	71.1	47.1	79.6	21.0	15.5	20.7	91.0	15.8	2.4	25.1	100	100	13.2	0.74	1.10	20.4
CG12-111	7.8	86.0	59.4	79.5	21.2	15.9	20.7	89.9	14.4	2.4	26.1	100	99	12.2	-0.19	-0.25	36.8
CG12-104	6.3	77.1	58.5	78.3	21.6	15.2	20.1	91.1	14.9	2.4	23.4	100	97	11.9	-1.02	-1.16	25.1
CG12-532007	8.8	72.3	58.4	79.6	21.5	16.4	21.7	92.3	13.1	2.4	27.7	100	100	12.1	2.64	2.90	27.2
CG12-106	4.8	86.5	58.8	79.3	21.0	14.6	19.7	89.9	12.5	2.4	19.5	100	99	14.0	0.36	0.54	11.7
CG12-101	7.1	77.6	58.8	79.4	22.4	15.4	21.0	90.8	15.5	2.4	24.3	100	85	13.0	1.36	1.51	23.4
CG12-108	7.1	86.5	58.8	79.3	20.4	15.0	20.1	90.0	16.0	2.4	23.9	100	100	13.7	1.08	1.31	22.4
CG12-060016	5.3	69.8	52.0	79.3	20.1	15.1	19.5	92.2	15.0	2.4	22.9	100	73	12.1	-0.71	-0.68	15.7
CP88-1165	7.1	58.1	58.8	79.3	22.0	15.5	20.1	93.8	14.5	2.4	26.7	90	100	10.6	0.78	0.77	21.2
CG12-018003	5.7	77.1	59.4	78.6	22.2	15.0	20.1	90.6	14.9	2.4	22.2	100	97	11.3	-0.42	-0.26	14.3
CG12-339052	6.4	74.0	58.8	77.7	20.6	14.7	20.3	89.0	17.0	2.4	22.0	100	71	12.2	2.45	2.50	21.3
CG12-311020	8.5	67.7	57.8	78.9	21.0	15.2	20.8	88.7	13.0	2.4	26.8	100	100	12.5	1.19	1.05	24.0
CG12-318005	6.4	82.9	58.5	79.5	20.4	13.9	19.3	88.5	16.6	2.4	21.1	100	86	11.7	0.86	0.91	21.2
CG12-606129	8.4	85.6	55.7	74.1	21.1	15.2	20.4	89.2	13.5	2.5	26.1	100	78	12.9	-0.33	-0.22	20.0
CP72-2086	8.9	80.6	57.8	79.1	21.8	16.9	22.2	92.8	16.4	2.5	26.9	95	100	12.0	1.02	0.92	28.6
CG12-620015	7.1	42.7	44.0	78.9	21.1	14.6	19.7	89.5	15.3	2.4	25.7	100	100	11.7	-1.03	-1.02	21.8
CG12-102	7.3	85.6	57.7	78.4	20.6	14.7	20.2	88.6	14.3	2.4	25.3	100	78	12.3	-1.64	-1.71	19.1
CG12-330012	6.6	74.1	57.5	77.6	22.7	16.4	21.7	91.4	12.0	2.5	25.3	100	44	12.4	-1.15	-0.62	17.5
CG12-030014	8.7	75.7	57.4	76.9	22.8	16.1	21.5	91.2	12.4	2.5	23.5	100	100	13.5	-0.91	-1.05	21.1
CG12-112	6.6	73.1	56.5	79.6	22.1	15.3	20.6	89.9	14.0	2.5	23.3	100	82	11.6	-0.44	-0.49	21.6
CG12-106001	9.0	80.9	58.0	79.3	22.7	16.1	21.0	92.9	15.0	2.6	27.9	100	100	12.7	-0.24	-0.52	29.0
CG12-005031	8.6	81.0	58.2	79.2	21.0	15.5	20.7	91.0	14.9	2.6	25.4	100	100	11.1	-0.36	-0.50	20.8
CG12-110	6.8	83.5	56.9	79.2	20.7	15.5	20.7	91.5	16.9	2.6	22.9	100	99	13.7	1.08	1.17	22.0
CG12-526031	9.7	75.5	51.6	79.3	22.8	16.9	22.1	92.2	12.7	2.6	26.6	100	96	11.3	-1.21	-1.39	22.8
CG12-276003	8.3	83.8	56.2	79.1	20.7	15.4	20.5	88.5	12.1	2.5	27.1	100	93	11.9	-1.35	-1.68	14.7
CG12-103	7.5	79.5	46.7	79.3	21.8	15.0	20.2	89.9	9.5	2.6	23.0	100	82	15.0	-1.82	-2.32	14.3
CG12-116	8.9	70.3	54.6	79.3	22.0	17.1	22.5	92.4	16.5	2.7	24.7	100	88	11.4	0.81	1.12	33.8
CG12-105	7.0	83.2	50.8	65.6	21.3	14.5	20.0	87.6	16.4	2.6	24.0	100	99	11.3	-2.10	-2.38	22.7
SP79-1287	9.6	75.1	58.4	79.4	20.5	15.5	20.9	90.3	14.8	2.6	27.7	95	100	11.7	0.76	0.48	29.2
CG12-114	6.8	86.5	58.8	81.0	22.5	15.3	20.8	89.6	16.7	2.6	22.0	100	100.0	12.1	-0.53	-0.58	22.0
CG12-326030	7.8	79.9	59.1	79.1	21.2	15.5	20.9	90.3	15.6	2.7	23.0	100	82.3	12.9	-1.27	-1.49	28.3
CG12-339025	8.6	80.3	58.1	79.5	22.2	14.6	19.9	90.5	14.8	2.7	22.4	100	100.0	11.9	-0.02	0.14	19.0
CG12-311019	9.5	76.9	52.5	73.9	22.5	15.2	21.0	89.2	13.9	2.7	27.0	100	99.3	11.4	0.56	0.33	29.4
CG12-115	8.3	77.1	58.8	79.3	19.6	13.0	17.7	89.4	18.5	2.7	28.3	100	100.0	12.6	-1.03	-1.29	25.8
CG12-324094	8.5	52.3	49.4	79.1	21.4	15.3	20.3	90.1	15.8	2.7	26.3	100	95.5	10.3	0.27	0.13	32.7
CG12-109	7.8	81.3	55.7	79.5	20.3	15.2	20.7	89.3	15.6	2.7	24.8	100	93.3	13.3	-1.68	-1.95	27.3
SP71-6161	8.3	86.5	58.8	79.3	21.3	16.3	21.7	91.2	16.0	2.8	23.0	100	100.0	11.4	-0.47	-0.29	29.2

brixprom_1, Brix medido en el campo; peso5stalk_2, Peso de una muestra de cinco tallos de caña; oqueleve_3, Oquedad leve; oqueinter_4, Oquedad intermedia; oqueve_5, Oquedad severa; pol_6, Pol%Caña (medida de la concentración del azúcar en la caña); brixjugo_7, Brix del jugo medido en el laboratorio; purity_8, Pureza del jugo, determinada en el laboratorio; pop_9, Población de tallos; alt_10, Altura de corte de la caña; diam_11, Diámetro medio de la caña; flower_12, Proporción de tallos florecidos; pith_13, Corcho o tejido interno seco en las cañas; fiber_14, Materia Seca de la muestra; TAH, t de azúcar por ha.

Tabla 4.50. Resumen de las Salidas de RINDSEL para los índices ESIM y RESIM para el conjunto de datos medio_no_flor_1ª_soca.

Índice	Estadístico	Valor del estadístico	Fracción superior (10% Variedades Seleccionadas)	Valor del índice
ESIM	Covarianza entre el índice y el VMG	0.2583723	CG12-116	2.642
	Varianza del índice	0.7499771	SP71-6161 (testigo)	2.450
	Varianza del VMG	0.04091631	CG12-526031	2.232
	Correlación entre el índice y el VMG	0.9999	CP72-2086 (testigo)	1.947
			CG12-106001	1.362
			CG12-109	1.188
RESIM	Covarianza entre el índice y el VMG	0.4455894	CG12-116	2.902
	Varianza del índice	1.208821	CG12-526031	2.599
	Varianza del VMG	0.09363591	SP71-6161 (testigo)	2.501
	Correlación entre el índice y el VMG	0.9999	CP72-2086 (testigo)	2.300
			CG12-106001	1.514
			CG12-311019	1.312

VMG: VMG: Valor de la Mejora Genética (*Breeding value*).

En general, las selecciones en ambos índices fueron muy similares y el orden de éstas varió poco (Tabla 4.50). Se apreció además que, en este ambiente de evaluación, los testigos se presentaron también con buenos desempeños, llegando a competir con los principales materiales genéticos. En este grupo destacó la variedad CG12-116 que fue bien calificada en los valores de ambos índices. Debido a que la correlación entre estos dos índices fue alta ($r=0.99$), es importante señalar que las variables que restringen el cálculo en **RESIM** fueron las que presentaron los valores más altos de ponderación en **ESIM**. Por tanto, **RESIM** fue equivalente a un **ESIM** donde solamente se consideraron tales variables. En este sentido, pude acotarse aún más la valoración de los genotipos en torno a estas variables (Promedio de °Brix en el campo, Peso de cinco tallos, Pol%caña, corcho y fibra). Por último, estas variables parece que tuvieron suficiente peso como para poder calificar apropiadamente a los genotipos de caña de azúcar en evaluación.

4.13.3.4 CPs como Índices de Desempeño para el conjunto de datos litoral_no_flor_plantia

Las funciones lineales desarrolladas como producto de un ACP se utilizaron como índices de desempeño, en particular, las tres primeras CPs. Al igual que en los dos conjuntos de datos analizados anteriormente, la CP1 estuvo asociada a las variables de concentración de

sacarosa (21.6% del total de la variabilidad), la CP2 estuvo altamente asociada a los caracteres de acumulación de biomasa (**peso5stalk_2**, **alt_10**, **diam_11** y **fiber_14**) (14.6% de la variabilidad) y la CP3 estuvo asociada a los caracteres de calidad del Tallo (Oquedad, Corcho, y Población). En su conjunto, estos tres componentes acumularon el 49.6% de la variabilidad (Tabla 4.14).

Los valores de heredabilidad de cada componente fueron relativamente elevados, disminuyendo de la CP1 a la CP3 (Tabla 4.51), lo cual está relacionado con la importancia de los componentes, así como su importancia en términos del valor genético de los materiales en estudio.

Tabla 4.51. Resumen de indicadores para los índices CP1, CP2 y CP3 para el conjunto de datos litoral_no_flor_plantía.

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY)	Varianza ambiental (RESIDUOS)	H ²	Valor estimado de Mejora genética = H ² *(ENTRY + RESIDUOS)
CP1	0	1.99554	0.834	0.827	2.340
CP2	0	1.20960	0.835	0.743	1.519
CP3	0	0.42830	0.836	0.506	0.640

H²= heredabilidad en el sentido amplio

Debido a que los signos de las cargas para cada variable de mayor importancia en las CP1 y CP2 son negativas (Tabla 4.15 y Figura 4.21), los materiales con los mejores desempeños fueron aquellos cuyos *scores* fueron los más bajos para la CP1 y la CP2. En el caso de la CP3, las cargas en las variables de mayor importancia fueron positivas, por lo que los genotipos con los mejores desempeños fueron aquellos con los *scores* más altos.

A diferencia de los dos conjuntos de datos anteriores, en este caso se encontraron algunas coincidencias en la valoración de los genotipos en estudio para los índices CP1, CP2 y CP3 (Tabla 4.52). Esto puede deberse a que una buena manifestación fenotípica se asoció a un buen nivel de calificación del desempeño según las CPs, y en la mayoría de variables de estudio pudo apreciarse este fenómeno (Tabla 4.49), de tal manera que cada una de ellas aportó un peso significativo, que a su vez se expresó en los índices de desempeño, en este caso las CPs. Para darle más sustento a este procedimiento, se realizó un análisis de varianza para un modelo lineal mixto con respecto a la CP1, el cual presentó un buen nivel de ajuste de

la variable en términos normales (Figura 4.37). Por otro lado, el análisis indicó que se presentaron diferencias significativas entre las medias de los genotipos en evaluación ($F=5.314$, ***) (Tabla 4.17).

Tabla 4.52. Resumen de scores para los índices CP 1, 2 y 3 para el conjunto de datos litoral_no_flor_plantía del ~25% de la fracción superior.

Genotipo	Score CP1	Genotipo	Score CP2	Genotipo	Score CP3
CG12-526031	-2.934	CG12-324094	-2.616	CG12-114	2.060
CG12-113	-2.686	SP79-1287	-2.542	CG12-110	1.838
CP72-2086 (testigo)	-2.216	CG12-311020	-2.301	CG12-115	1.707
CG12-111	-2.139	CG12-105	-2.110	CG12-339025	1.314
CG12-532007	-2.017	CG12-620015	-1.676	CG12-620015	1.127
CG12-101	-1.830	CG12-109	-1.631	SP71-6161	1.117
CG12-104	-1.723	CP88-1165	-1.586	CG12-326030	1.088
CG12-116	-1.685	CG12-317012	-1.432	SP79-1287	1.065
CG12-005031	-1.646	CG12-311019	-1.261	CG12-311019	0.731
CG12-106001	-1.455	CG12-606129	-1.138	CG12-111	0.611
	-2.934	CG12-115	-1.019	CP88-1165	0.587

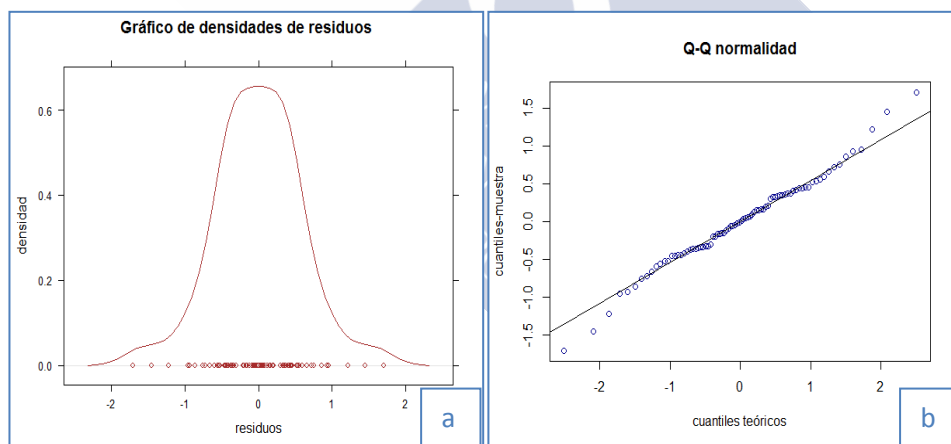


Figura 4.37. Gráficos de densidades (a) y cuantiles (b) para la variable CP1 del conjunto de datos litoral_no_flor_plantía.

Se encontró un grupo de genotipos que presentaron un buen nivel de diferenciación por pares con el resto (Figura 4.26), entre los cuales destacaron las variedades CG12-606129, CG12-115, CG12-110, CG12-276003, CG12-339025 y CG12-018003; con un buen desempeño en términos de las variables relacionadas con la sacarosa (CP1), superando a los testigos propuestos.

De acuerdo con el esquema de selección anidada, los genotipos que fueron seleccionados de acuerdo a los *scores* de las CPs fueron CG12-311019, CG12-111, CG12-106001, y CG12-116 (Tabla 4.53). Estas selecciones no coincidieron con las selecciones hechas con los cuatro índices discutidos con anterioridad. Si bien se presentó una coincidencia en la selección, ésta debería ser considerada aleatoria. En el presente caso, debido a que las “cargas” (*eigenvectors*) fueron de orden negativo para las CP1 y CP2, los genotipos con los *scores* más bajos fueron los de mejor desempeño. Lo contrario ocurrió para la CP3, donde las “cargas” más altas fueron positivas, en consecuencia, el mejor desempeño lo determinaron los *scores* más altos.

Tabla 4.53. Fracción superior del 25% (10 variedades) seleccionada por la CP1. Comparación por pares ajustada por el método de Tukey y ordenación por las CP2 y CP3 para la selección final (~10% de la fracción superior, en verde).

Genotipo	Tukey CP1	Score CP1	Genotipo	Score CP2	Genotipo	Score CP3
CG12-526031	E	-2.934	CG12-324094	-2.616	CG12-311019	0.731
CG12-113	CE	-2.685	CG12-311019	-1.261	CG12-111	0.611
CG12-111	CDE	-2.139	CG12-116	-0.620	CG12-106001	0.434
CG12-532007	ABCDE	-2.017	CG12-106001	-0.613	CG12-116	0.310
CG12-101	ABCDE	-1.830	CG12-526031	-0.272		
CG12-104	ABCDE	-1.723	CG12-111	-0.236		
CG12-116	ABCDE	-1.685				
CG12-005031	ABCDE	-1.646				
CG12-106001	ABCDEF	-1.455				
CG12-311019	ABCDE	-1.430				
CG12-324094	ABCDEF	-0.893				
CP72-2086	BCDE	-2.216				
CP88-1165	ABCDEF	-1.359				

4.13.3.5 TAH (t De Azúcar Por ha) y BLUP (*Best Linear Unbiased Prediction*)

La variable TAH para este conjunto de datos presentó un valor fenotípico medio de 22.85 (Tabla 4.2). La heredabilidad para TAH fue 0.702, y el valor estimado de mejora genética (VMG) 27.57 (Tabla 4.54). En este grupo de datos, el desempeño general fue superior y se presentó como uno de los mejores entre todos los conjuntos de datos estudiados. Esto es producto, como ya se mencionó antes, del mejor desarrollo general de las plantas y a su buen desempeño fenotípico, el cual fue liderado principalmente por las variables relacionadas con

la acumulación de biomasa. Se sigue que un buen desarrollo fenotípico se reflejó en un mejor nivel de diferenciación de los genotipos, y en un mejor discernimiento del valor genético de los materiales en evaluación.

Tabla 4.54. Resumen de indicadores de los índices TAH para el conjunto de datos litoral_no_flor_plantía

Índice	Valor fenotípico medio	Varianza genotípica (ENTRY)	Varianza ambiental (RESIDUOS)	H ²	Valor de Mejora genética (VMG)=H ² *(ENTR Y + RESIDUOS)
TAH	22.85 ± 6.5	21.269	18.011	0.702	27.574

H²= heredabilidad en el sentido amplio

Al emplear solamente el TAH como un índice de desempeño se obtuvo, después de un análisis de variación para la variable TAH en el esquema de un modelo lineal mixto, que las diferencias de las medias de TAH entre los genotipos evaluados fueron altamente significativas (F= 3.1501***).

Si bien los ajustes de BLUP no fueron en importantes en magnitud, sí se evidenció que, comparado con otros conjuntos de datos, hubo una mayor frecuencia de valores que debieron ser ajustados (Figura 4.38.a). En la distribución de frecuencias absolutas de los BLUPs (Figura 4.38.b) se encontró que las frecuencias absolutas de valores altos de BLUP (>0) fueron relativamente bajas, en tanto que los valores por debajo de cero fueron las que presentaron las frecuencias más elevadas. Esto significa que, en general, la mayoría de genotipos presentaron valores de desempeño por debajo de la media y, por tanto, la mayoría de materiales en estudio no fueron aptos para la selección, esto a pesar de que los valores del TAH para este conjunto de datos se consideraron altos. Esto significa que, dentro de este grupo, los mejores genotipos presentaron desempeños realmente altos (Tabla 4.2).

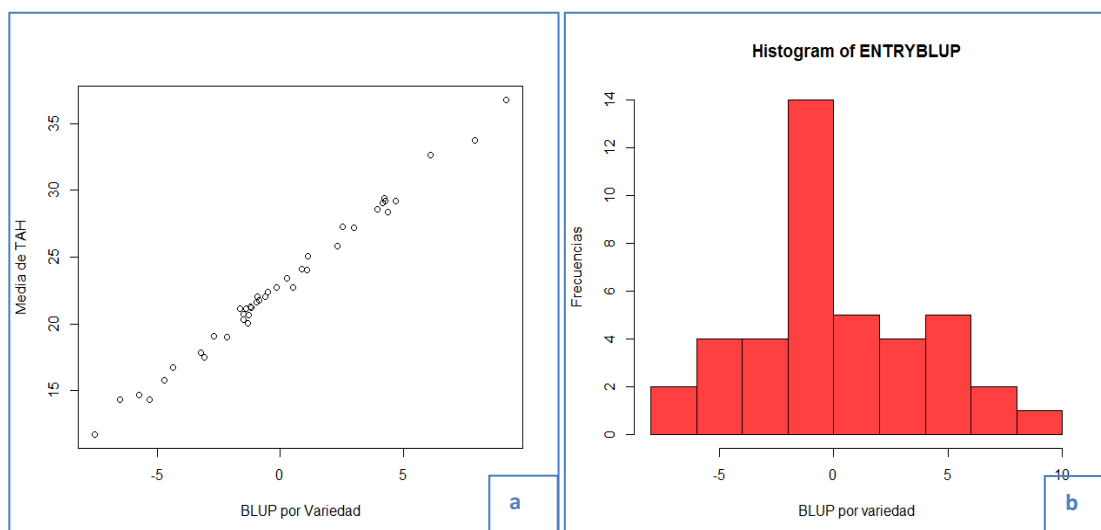


Figura 4.38. Ajuste de BLUP con respecto de la media de TAH (a) y distribución de las frecuencias de BLUP (b).

La selección del 10% de los genotipos con respecto al TAH incluyeron un grupo de variedades CGs que, en general, mostraron un buen desempeño con respecto a esta variable (Tabla 4.36), ya que los valores del TAH fueron superiores a 30. Destacó la variedad CG12-111 con un alto valor de TAH (36.78). Esta variedad, en particular, presentó el valor más alto en TAH comparada con cualquier otra variedad en cualquiera de los siete conjuntos de datos. En todo caso, el número de diferencias por pares fue relativamente bajo, y fue únicamente la variedad CG12-111 la que se diferenció de otras cinco variedades; así como la variedad CG12-116, la cual se distinguió estadísticamente de solamente una variedad (CG12-606129) (Figura 4.26).

Según los valores de TAH para la fracción superior (25%) (Tabla 4.55), los genotipos seleccionados fueron CG12-111, CG12-116, CG12-324094, y CG12-311019. Según los valores de BLUP para cada genotipo, se observó que el orden varió; en esta nueva ordenación los genotipos sugeridos para ser seleccionados fueron CG12-111, CG12-116, CG12-324094, y CG12-326030. Al realizar la corrección de los datos en términos de BLUP se apreció que la selección varió en al menos un genotipo.

Tabla 4.55. Ordenamiento descendente de la fracción superior (~25%) de genotipos según los valores de TAH y BLUP. Separación de medias ajustada por la metodología de Tukey.

Genotipo	TAH	Tukey	Genotipo	BLUP
CG12-111	36.78	ABCDEF	CG12-111	9.205
CG12-116	33.75	ABE	CG12-116	7.920
CG12-324094	32.65	ABCDEF	CG12-324094	6.100
CG12-311019	29.37	ABCDEF	CG12-326030	4.381
SP79-1287 (testigo)	29.19	ABCDEF	SP71-6161(testigo)	4.699
SP71-6161(testigo)	29.16	ABCDF	SP79-1287 (testigo)	4.279
CG12-106001	29.02	ABCDEF	CG12-311019	4.244
CP72-2086 (testigo)	28.56	ABCDEF	CG12-106001	4.163
CG12-326030	28.34	ABCDF	CP72-2086 (testigo)	3.966
CG12-109	27.29	ABCDEF	CG12-109	2.542

4.13.3.6 Regresión Logit para establecer relaciones entre índices y la base de comparación de selección por competencia para el conjunto de datos litoral_no_flor_plantía

Las funciones de regresión Logit fueron más claras para los índices TAH y BLUP (Tabla 4.56, Figura 4.39), en tanto que para cuatro del resto de las funciones que se presentaron en la Figura, la relación fue baja. Esto se relaciona con el hecho de que el TAH es un argumento importante para el fitomejorador a fin de decidir la selección. Así, en este grupo de datos se comprobó que, según los requerimientos por competencia para los genotipos en evaluación, el TAH y el BLUP del TAH fueron los valores que mejor se ajustaron.

Tabla 4.56. Valores De Pseudo R² de Mcfadden de una regresión Logit para el modelo general $\ln[p/(1-p)] = a + bx + e$ para el conjunto de datos medio_no_flor_1ª_soca.

índice	Modelo	Prueba de Significación Pr(>x ²)	Pseudo-r ² con la base selección por competencia	Eficiencia con respecto al r ² más alto
TAH	$\ln[Si/(1-Si)] = -17.5663 + 1.0879 \text{ TAH}$	<0.0001	0.724	1
BLUP	$\ln[Si/(1-Si)] = 6.6865 + 1.4725 \text{ BLUP}$	<0.0001	0.719	0.99
RESIM	$\ln[Si/(1-Si)] = 2.5342 + 1.0150 \text{ RESIM}$	0.025	0.166	0.23
CP2	$\ln[Si/(1-Si)] = 2.5302 - 1.0765 \text{ CP2}$	0.035	0.145	0.20
ESIM	$\ln[Si/(1-Si)] = 2.4762 + 1.0662 \text{ ESIM}$	0.0369	0.145	0.20
CP1	$\ln[Si/(1-Si)] = 2.3611 - 0.6851 \text{ CP1}$	0.056	0.120	0.17
KN	$\ln[Si/(1-Si)] = 2.1004 + 1.8783 \text{ KN}$	0.249	0.044	0.06
Smith	$\ln[Si/(1-Si)] = 2.0467 - 0.5122 \text{ Smith}$	0.353	0.028	0.04
CP3	$\ln[Si/(1-Si)] = 2.0122 + 0.2972 \text{ CP3}$	0.492	0.016	0.02

a=intercepto, b=coeficiente, X=Índice, e=término del error, p=probabilidad de ocurrencia de "sí", 1-p=probabilidad de ocurrencia de "no".

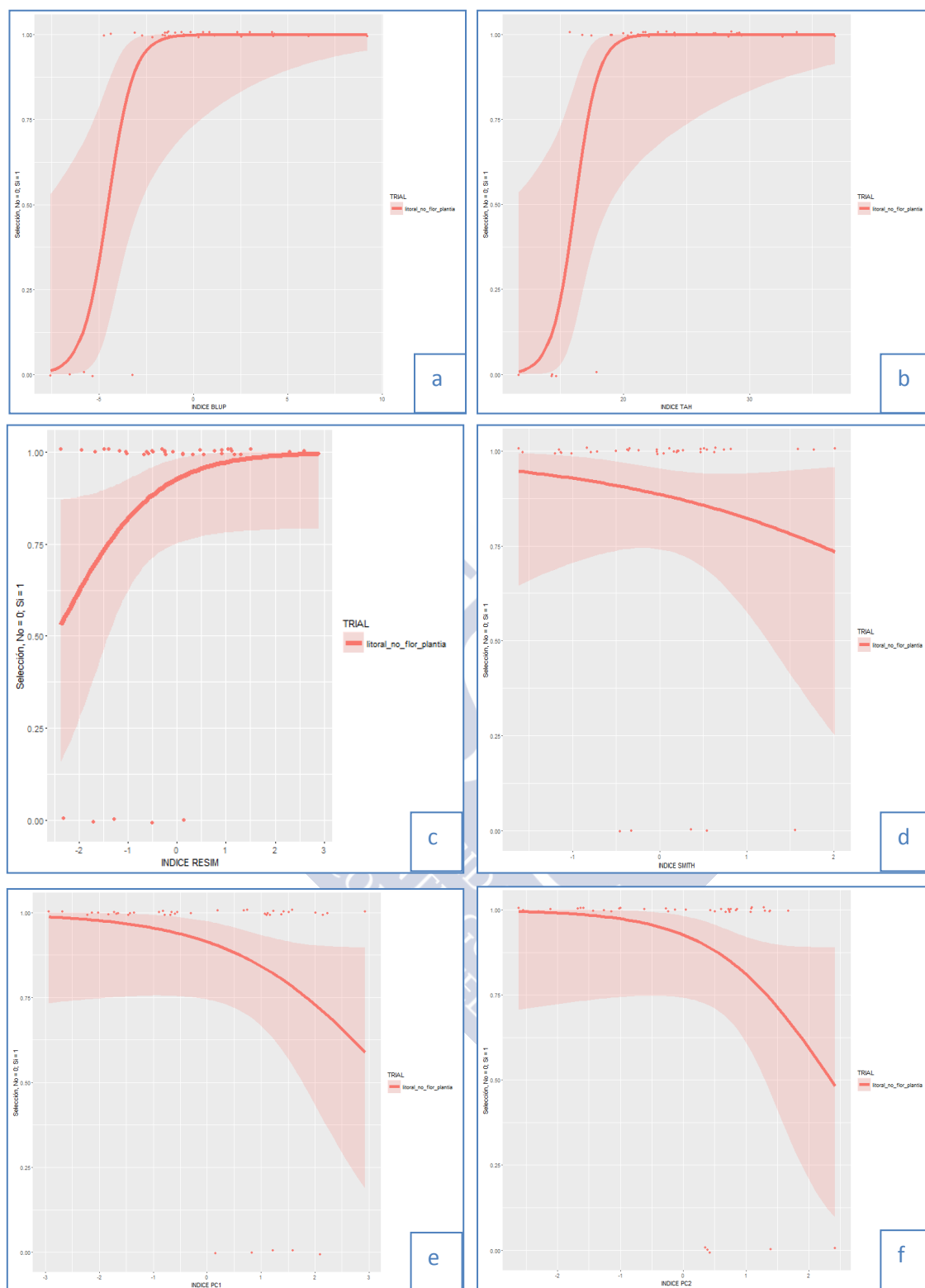


Figura 4.39. Gráficos de las funciones **Logit** para los índices: a) BLUP, b) TAH, c) RESIM, d) SMITH, e) PC1, Y f) PC2.

4.13.4 Las regresiones Logit en términos generales

En la Figura 4.40 se presentan varios gráficos sobre las funciones de regresión entre los diferentes índices y la base de comparación, la competencia del fitomejorador, según el estrato en estudio o el patrón de floración. En esta Figura se resume los patrones de regresión planteados por cada uno de los tres conjuntos de datos. Se observa en la Figura 4.40 (a, b, c, d) que para TAH y BLUP, el ajuste fue alto para todos los casos, ya sea según el patrón de floración o según el estrato de evaluación. Por otro lado, se aprecia que el índice ESIM, presentado aquí como referencia, no mostró un nivel apropiado de valoración según los estándares de la competencia de los fitomejoradores, por lo que se sigue que no es apropiado. En resumen, fue TAH y BLUP del TAH las estimaciones que mejor “indicaron” (estimaron) el valor genético de los materiales de caña de azúcar en evaluación.



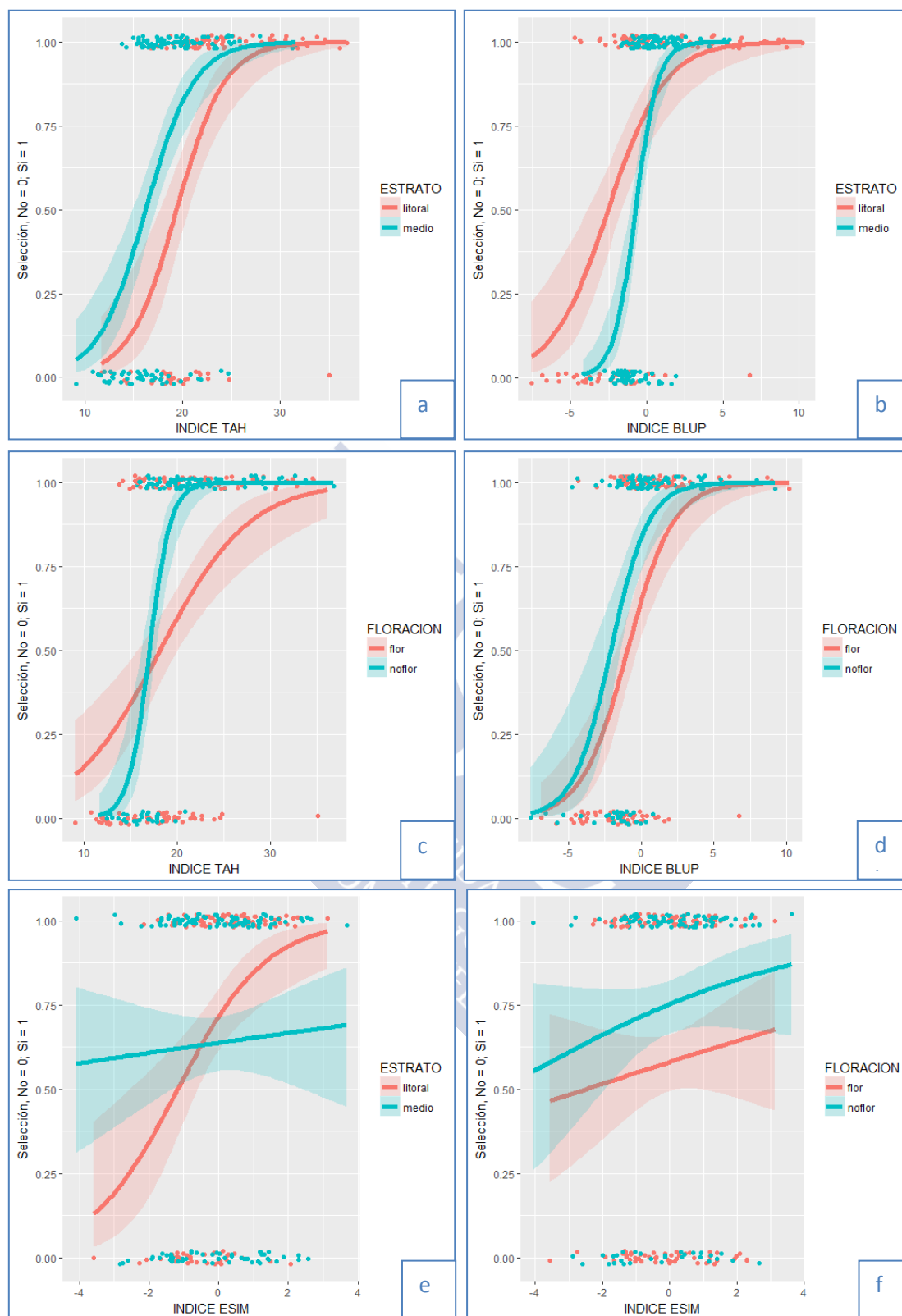


Figura 4.40. Gráficos de Regresión lineal Logit para la relación entre la base de comparación de selección por competencia (“Si” o “No”) con los índices TAH y BLUP según el estrato en estudio (a y b) o el patrón de floración (c y d) y ESIM según estrato de evaluación (e) o el patrón de floración (f).





5. DISCUSIÓN



5 DISCUSIÓN

5.1 SOBRE LOS RESULTADOS GENERALES

5.1.1 Medidas estadísticas generales

Algunos valores del coeficiente de variación de los valores de medidas estadísticas básicas fueron relativamente elevados, como sucedió para la variable **peso5stalk_2** (Tabla 4.1). En nuestros resultados, la diferencia global entre promedios de **peso5stalk_2** según los estratos de crecimiento (0.67 kg) fue altamente significativa ($t=4.42$, ****, Tabla 4.3). Por lo tanto, se deduce que las condiciones ambientales (Tabla 3.1 en la sección de Materiales y Métodos) influyeron sobre esta variable, la cual, en consecuencia, se presentó como un carácter de variación cuantitativa (Bos y Caligari, 2007). Shanti *et al.*, (2013) encontraron que la ganancia genética con respecto al carácter de peso de los tallos fue relativamente baja, al igual que la eficiencia de selección directa, comparada ésta con la selección hecha con un índice de selección. El resto de variables se mostraron relativamente estables, y los errores estándar mostraron esa tendencia. Para la variable **peso5stalk_2**, los valores superiores a 7 kilos presentaron, aproximadamente, productividades de más de 130 t por ha, , con una elevada frecuencia en ese rango de valores en nuestros resultados (Figura 4.1), lo que concuerda con otros resultados en estudios similares (Rosales-Longo *et al.*, 2016b).

Los valores superiores al 15% de Pol % caña (**pol_6**) se consideran como un valor bueno (Orozco *et al.*, 2012, Berding *et al.*, 2004; Ming *et al.*, 2006). Larrahondo (1995) indica que la concentración de azúcar medida en °Brix (**brixjugo_7**) puede variar entre 10 y 16. En nuestros resultados, estos valores se alcanzaron como promedio de 14.90% de todos los experimentos (Tabla 4.1 y Figura 4.4), debido a que se contó con una variación alta en términos de las frecuencias de genotipos con bajos valores de **pol_6** (Min=7.83) y frecuencias de genotipos con altos valores de **pol_6** (Max=18.30). Khan *et al.* (2012) encontraron valores comparables

de contenido de sacarosa entre 12.08% y 16.45%, en tanto que Tena *et al.* (2016a) encontraron valores que varían entre 12.56% y 14.48%. Lo anterior indica que en el presente trabajo se encontró una mayor variabilidad y que una fracción de esta es explotable en cuanto a este carácter.

Las condiciones de valores menores en temperatura, menor luminosidad y mayor precipitación, en el estrato medio con respecto al estrato litoral, propiciaron la mayor acumulación de ciertos componentes de la caña de azúcar, tales como la sacarosa y la fibra. A menores temperaturas las tasas fotosintéticas disminuyen como consecuencia de una actividad metabólica reducida, lo que deriva en una menor cantidad de azúcares del culmo para el gasto energético (Watt *et al.*, 2014). Debido a que la caña de azúcar es una planta C4 (Sage *et al.*, 2014), ésta reacciona ante la mayor radiación recibida en el estrato litoral, lo cual propició una mayor acumulación de biomasa, comparada ésta con la que se midió en el estrato medio.

Con respecto a la variable fibra (**fiber_14**), la menor diferencia de los promedios según el estrato (Tabla 4.3), significativa al 5% (*), indicó que el ambiente ejerció menos sobre esta variable, con una tendencia a favor de los genotipos del estrato medio. Por otro lado, siempre en el caso de la variable **fiber_14**, la diferencia fue altamente significativa entre los promedios, según el patrón de floración ($t=6.83$, *****) (Tabla 4.3), lo cual es un indicativo de que el componente genético ejerció un fuerte impacto sobre esta variable. En general, los genotipos no florecedores tienden a contener más materia seca que los genotipos florecedores. La presencia de flor (**flower_12**) está altamente relacionada con la de corcho (**pith_13**). Esto es, a medida que los entrenudos superiores se elongan, las paredes del parénquima de almacenamiento en el centro del culmo cesan el desarrollo de paredes celulares y pueden, dependiendo del clima y del clon, colapsar hasta formar un tejido esponjoso (corcho) con cavidades de aire que pueden coalescer y, finalmente, producir oquedad (Moore y Berding, 2014).

Al observar los resultados sobre la concentración de sacarosa (**pol_6**) (Tabla 4.3) se advierte que el ambiente ejerció una fuerte influencia sobre la variación de este carácter, lo cual, en general, es consistente con lo que se ha encontrado en otros conjuntos de experimentos, en donde se observa que la concentración de sacarosa es superior en aquellos experimentos que se realizaron en el estrato medio, comparados con los valores de los estratos bajo y litoral de la zona productiva de la agroindustria azucarera guatemalteca (Rosales-

Longo, 2016b). Como se ha indicado anteriormente, la acumulación de sacarosa en los culmos se reduce en zonas con menores temperaturas, como consecuencia de una actividad metabólica reducida (Watt *et al.*, 2014; Sage *et al.*, 2014).

Con respecto a lo que se observó para **pol_6**, también es válido para las variables °Brix en el campo (**brixprom_1**) y el °Brix medido en el laboratorio (**brixjugo_7**), lo cual también es consistente con lo que se han encontrado en otros experimentos de evaluación de genotipos de caña de azúcar en los Estados II y III del programa de mejoramiento genético de CENGICANÑA en Guatemala (Rosales-Longo, *et al.*, 2016a y 2016b). En general, una reducción en la acumulación de biomasa se ve reflejada finalmente en un aumento en la concentración de sacarosa. Los cambios en la temperatura y la disponibilidad de agua pueden reducir la elongación de los entrenudos, el crecimiento del dosel de las plantas y, a la vez, como consecuencia, aumenta la proporción de sacarosa (Inman-Bamber *et al.*, 2011).

En todo caso, pueden obtenerse respuestas genéticas con respecto al contenido de sacarosa según las diferencias que se encontraron en este estudio. Inman-Bamber *et al.* (2011) indicaron que las ganancias pueden estudiarse desde el punto de vista genético/ fisiológico en dos formas: a) la proporción de la varianza atribuible a los genotipos con respecto a este carácter, tal como se realizó en este estudio, la cual es debida a la variación anatómica en las células de almacenamiento; y b) las diferentes respuestas de los genotipos ante las condiciones ambientales y de maduración, lo cual también se estudió en este trabajo.

Con respecto a las tendencias observadas en la variable TAH, resultados en términos de diferencias por estrato y por patrón de floración, éstas fueron consistentes con lo que usualmente se encuentra en la producción comercial de azúcar en Guatemala (CENGICANÑA, 2011-2017). En este sentido es posible indicar que las condiciones de cultivo en el estrato litoral favorecieron una mayor productividad; sin embargo, la variación en los valores de la producción fue mayor, comparada con lo que se apreció en el estrato medio. Es decir, existen mayores posibilidades de que se alcancen bajas producciones en el estrato litoral, de forma análoga a lo que ocurre en el estrato medio, pero las más altas producciones que se observaron en el estrato litoral no pudieron ser alcanzadas en el estrato medio (Tablas 4.3 y 4.4), debido a la interacción del genotipo con el ambiente (Bos y Caligari, 2007), ya que la variable está construida a partir de variables de variación cuantitativa (biomasa) y del contenido de

sacarosa, la cual posee valores de heredabilidad relativamente altos (Jackson, 2005; Inman-Bamber *et al.*, 2011), extremo que también se verificó en este trabajo.

5.1.2 Análisis de Varianza

Con respecto a los datos meteorológicos de precipitación, temperatura y radiación solar, se derivó que mientras la precipitación se reduce según el avance de la zafra, la radiación aumenta y esto favorece una menor disponibilidad de agua para las plantas, un resumen de las condiciones meteorológicas se presentan en la Tabla 3.1.; lo anterior si no se cuenta con el riego correspondiente. Este fenómeno influye, en buena medida, en los valores de concentración de azúcar, ya que al haber menor disponibilidad de humedad el valor de concentración del azúcar aumenta, si bien el monto absoluto de azúcar no es el que necesariamente varía (Watt *et al.*, 2014; Inman-Bamber, 2014). Sin embargo, a pesar de la influencia del ambiente, existe un componente que indica que la mayor concentración de azúcar, reflejada tanto en las variables de medición del contenido de azúcar (Tabla 4.4) así como en el TAH (Floración = 6.36, **, Tabla 4.6), está influenciada genéticamente de manera importante. Lo anterior es consistente con lo encontrado en varios programas de mejora genética (Berding *et al.*, 2004; Jackson, 2005; Inman-Bamber *et al.*, 2011). Lo anterior es válido tanto para las comparaciones según el patrón de floración como para el estrato de evaluación.

Con respecto a lo anteriormente discutido, en este estudio, para las tres variables que miden este aspecto (**brixprom_1**, **brixjugo_7** y **pol_6**), se registró que los genotipos **no_flor** consistentemente tienen valores más altos que los que se presentan los genotipos **flor**. En consecuencia, se sigue que, si bien hay una influencia ambiental, hay un efecto de las diferencias como producto de la acción del componente genético (Tabla 4.3), lo cual se ve reflejado en los correspondientes valores de heredabilidad para las variables en mención (Tabla 4.7). Los valores altos de heredabilidad en este carácter son consistentes con los hallados en otros programas de mejora genética (Berding *et al.*, 2004; Jackson, 2005; Inman-Bamber *et al.*, 2011; Almeida *et al.*, 2014; Tena *et al.*, 2016b). Por otro lado, con respecto a lo anterior, es natural pensar que las varianzas relacionadas con el componente genético tienen una importancia relativa y que, para las variables relacionadas con la concentración de azúcar, estas varianzas reflejan los efectos sobre los valores de heredabilidad, por ejemplo para **pol_6**

(Tabla 4.7). Por tanto, se sigue que sí hay un efecto genético parcial relacionado con la mayor concentración de azúcar en las variedades **no_flor**, comparadas éstas con las variedades **flor**.

En términos de la fuente de variación genotipo (variedad o ENTRY), se encontraron para todas las variables diferencias estadísticas en menos del 1% (****). En consecuencia, se determinó que hay un efecto genético en la diferenciación de los promedios para todas las variables en estudio. Esto es consistente con estudios donde se evalúan materiales en diferentes condiciones (Jackson *et al.*, 1995; Khan *et al.*, 2012; Tena *et al.*, 2016b).

En un análisis combinado de varianza para la variable TAH, según se aprecia en la Tabla 4.6, existen diferencias significativas en cuanto a las medias de la mayoría de variables estudiadas. Estas diferencias son el reflejo de la naturaleza poligénica de la variable en cuestión, pero además muestran lo mucho que varían las respuestas de la caña de azúcar con respecto al ambiente. Como ya se apuntó antes, los genotipos de caña de azúcar, por su naturaleza poliploide y aneuploide, tienen respuestas altamente diferenciales con respecto al ambiente (Grivet *et al.*, 2004; Ming *et al.*, 2006). Por lo observado, las variaciones climáticas, de suelo y manejo agronómico influyen notablemente en el desempeño de la caña de azúcar, tal y como también lo plantean Berding *et al.* (2004); Gilbert *et al.* (2004) y Ming *et al.* (2006).

Para todas las variables estudiadas (14) en todos los análisis se obtuvo que, para la fuente de variación **Variedad** (ENTRY) (Tabla 4.4), las diferencias en los promedios fueron altamente significativas, por lo que se sigue que las diferencias en los promedios de todas las variables estudiadas fueron funciones de los materiales genéticos. Otros autores también reportan resultados en los cuales, prácticamente todas las variables estudiadas presentan diferencias estadísticas para la fuente de variación genotipo (Ftwi *et al.*, 2016; Tena *et al.*, 2016a,b; Khan *et al.*, 2012), aunque las variables en evaluación hayan sido relativamente diferentes en estos estudios.

Con respecto a la fuente de variación **Ciclo** (Tabla 4.4), las diferencias estadísticas encontradas pueden ser atribuidas a un efecto puramente genético, puesto que los genotipos evaluados para cada ciclo en cada experimento, fueron diferentes, excepto para los cultivares “testigo”, los cuales generaron el error experimental necesario para la elaboración de los análisis de variación combinados.

5.2 CORRELACIONES

En general, se considera que los valores de altura, población y diámetro son componentes de las variables de acumulación de biomasa (Berding *et al.*, 2004; Ming *et al.*, 2006; Orozco *et al.*, 2012). Sin embargo, en los resultados de este trabajo se observó que la influencia de la población (**pop_9**) sobre acumulación de biomasa es más bien baja ($r=-0.11$, *, Tabla 4.5). Las variables que más se asociaron a la acumulación de biomasa (**peso5stalk_2**) fueron el diámetro (**diam_11**) y la altura de corte (**alt_10**), por lo que pueden considerarse como las principales medidas indirectas o componentes de la misma. En general el diámetro y la altura son componentes del peso de un tallo en caña de azúcar (Ong'ala *et al.*, 2016; Tena *et al.*, 2016a,b; Viveros *et al.*, 2015).

Por otro lado, es razonable la correlación entre estas variables, si se entiende que una mayor altura y mayor diámetro supondrán mayor acumulación de biomasa en el tallo (Tabla 4.5), lo que es consistente con lo observado en variables de similar medición en otros trabajos (Kadian y Mehla, 2006; Viveros *et al.*, 2015; Khan *et al.*, 2012; Tena *et al.*, 2016b).

Las correlaciones entre variables asociadas con la concentración de azúcar fueron relativamente altas (Tabla 4.5), lo que supone un hallazgo útil, particularmente en el caso de cuando se desea que las medidas preliminares que se realizan en el campo del contenido de azúcar (**brixprom_1**) sean un indicador confiable preliminar sobre el desempeño de los genotipos en evaluación con respecto a este aspecto. En este sentido, es posible utilizar con un buen nivel de confianza, los valores de °Brix en el campo (**brixprom_1**) para ponderar inicialmente a los genotipos de caña, ya sea previo a la toma de una muestra para llevarla al laboratorio, o bien para ser utilizada directamente como un elemento para la selección (Figuras. 4.8 y 4.9). Los valores de estas correlaciones son relativamente comunes en otros trabajos y ayudan a caracterizar y estructurar las poblaciones de clones de caña de azúcar en evaluación (Khan *et al.*, 2012; Viveros *et al.*, 2015; Tena *et al.*, 2016a,b).

Los °Brix (**brixprom_1**) medidos directamente en el campo tienen importancia en el total del azúcar recuperable para un genotipo en particular y, en algunos casos, el efecto de ésta variable es más alto que el Pol % jugo, o su complemento, Pol % caña (**pol_6**) (Lombardi *et al.*, 2015; Barbosa *et al.*, 2017). La variable °Brix (**brixprom_1**) también ha sido utilizada para el cálculo de Toneladas °Brix ha⁻¹, un cercano equivalente a TAH, dicha variable también ha

sido utilizada para la derivación de los valores genéticos por medio de BLUP's (De Oliveira *et al.*, 2013), lo cual adiciona un valor importante al °Brix medido en el campo, en ausencia de datos de Pol % caña (**pol_6**).

Con respecto a las regresiones entre las variables °Brix en Jugo (**brixjugo_7**) vs °Brix en campo (**brixprom_1**) y Pol % Caña (**pol_6**) vs °Brix en campo (**brixprom_1**), al tomar en cuenta el estrato de evaluación, las tendencias entre las variables relacionadas con la concentración de azúcar (Figura 4.9) siguieron la misma tendencia que cuando se hizo el análisis considerando todos los datos (Figura 4.8). Otros trabajos han encontrado valores de correlación inferiores a los hallados en este trabajo, sin embargo, las tendencias generales fueron similares a las que aquí se discuten (Tena *et al.*, 2016a,b; Ong'ala *et al.*, 2016; Kadian y Mehla, 2006; Viveros *et al.*, 2015; Khan *et al.*, 2012).

No se encontró relación entre variables de biomasa con respecto a las variables de concentración de azúcar (Tabla 4.5). Esto indica que los componentes genéticos para cada conjunto de variables son de naturalezas diferentes y además son independientes. Estos fenómenos están asociados más bien con la partición del Carbono dentro de la planta, en términos de la formación de celulosa y hemicelulosas (relacionadas a su vez con la acumulación de biomasa) y en términos de la acumulación directa de sacarosa en los tallos, los cuales son dos caminos diferentes, y su estudio molecular indica que son procesos fisiológicos diferentes (Van der Merwe y Botha, 2014). En general, estos dos tipos de variables (biomasa vs contenido de azúcar) no guardan relación entre sí y no permiten la evaluación y selección simultánea, tal como lo plantea Jackson, (2005). En este sentido el mismo autor sugiere el estudio de las dinámicas genéticas (varianza genética aditiva, heredabilidad en el sentido estrecho, o correlaciones genéticas entre caracteres) con el fin de desarrollar índices óptimos de selección (Jackson, 2005). Por otro lado, la influencia del ambiente hace que, fisiológicamente, los asimilados se dividan en la planta y se utilicen en crecimiento y no en almacenamiento (Inman-Bamber *et al.*, 2011), lo que supone que debe haber un manejo agronómico especial para aquellos genotipos que presentan altos contenidos de sacarosa, como el empleo de maduradores (Inman-Bamber, 2014).

La correlación negativa entre Población de tallos (**pop_9**) y Pol % Caña (**pol_6**) (Figura 4.11) no es fácil de explicar, por cuanto no se encuentra un punto claro de relación entre cada

par de comparaciones, ya que éstas son de naturalezas distintas. En todo caso, para un genotipo en particular, una alta población de tallos por metro lineal, puede, en forma importante (valor de la correlación entre **pop_9** y **pol_6**, -0.45, ****; Tabla 4.5), resultar en una baja concentración de sacarosa. Las evaluaciones deben tomar en cuenta estos temas para buscar el equilibrio más apropiado en términos de valoración y selección.

En los ACP (Figura 4.13) también se encontró una tendencia similar con respecto a esta relación: los vectores relacionados con el azúcar tienen una dirección opuesta al vector de la variable **pop_9**, lo cual es razonable ya que éstos son derivados de la matriz de varianzas covarianzas entre variables. Este es un tema que debe seguir siendo evaluado puesto que, en términos generales, una buena variedad es aquella cuya población es mayor de 15 tallos por metro lineal (Berding *et al.*, 2004; Ming *et al.*, 2006; Orozco *et al.*, 2012). En este sentido debe buscarse encontrar un equilibrio que permita tener una variedad con una buena población, sin comprometer el contenido de azúcar.

Con respecto al tema del párrafo anterior, Watt *et al.* (2014) muestran que en caña de azúcar existe una alta relación entre los órganos fuente de azúcar (hojas) y los reservorios naturales en la especie (hojas y tallos). En experimentos relacionados con caña de azúcar (Watt *et al.*, 2014) se observó que cuanto más sombreados puedan estar los órganos fuente (hojas) menor es su actividad fotosintética y termina por convertirse en un órgano reservorio; y lo que es más importante, este fenómeno propicia que la acumulación de azúcares en el culmo se reduzca notablemente, en particular si el período de sombra es muy prolongado. Por lo tanto, es probable que una mayor cantidad de tallos por metro lineal induzca una mayor cantidad de biomasa y, en consecuencia, mayores niveles de sombra, por lo que las plantas tienden a acumular una buena cantidad de azúcar en las hojas bajas en vez de hacerlo en los culmos. Por otro lado, una arquitectura de planta diferente, como hojas erectas, también permitiría una mayor incidencia de luz, de tal manera que las hojas bajas mantengan su actividad fotosintética y permitan una mayor acumulación de azúcares en los culmos. En contraste con lo planteado al respecto, Ma *et al.* (2017) encontraron en maíz que una mayor densidad de plantas puede influir en el incremento de la concentración de azúcares en hasta un 25%. Por otro lado, Begna *et al.* (1997) encontraron, también en maíz, que se puede lograr un incremento en la producción de grano y otras variables al aumentar la densidad de cultivo

hasta 130,000 plantas ha⁻¹, como producto de utilizar híbridos de maíz de estatura reducida. Lo anterior hace entender que existe en las gramíneas un efecto genético que puede ser aprovechado en este sentido. Sin embargo, en caña de azúcar el largo de los tallos es un componente muy importante de la producción de biomasa en el campo, por lo que una estatura baja de los mismos no parece ser ventajosa (Tablas 4.3, 4.4 y 4.5). En todo caso, debe establecerse un procedimiento equilibrado de selección para permitir que el contenido de azúcar pueda ser extraído sin comprometer costos en corte, alce, transporte y molienda, como consecuencia del uso de genotipos con niveles muy altos de producción de biomasa. Debe establecerse un equilibrio, genético, anatómico y fisiológico en la selección de clones (Jackson, 2005; Inman-Bamber *et al.*, 2011), preferiblemente mediante el uso de índices óptimos de selección (Jackson, 2005).

Se observó que la mayoría de las variables en estudio guardan una importancia relativa con respecto a tres grandes grupos: acumulación de biomasa, concentración de azúcar y calidad interna de los tallos. Estas variables muestran que su utilidad como caracterizadores es considerable como se planteó anteriormente (Berding *et al.*, 2004; Manners, 2011). Con respecto a lo anterior Berding *et al.* (2004) indican que las varianzas genéticas aditivas son más importantes que las no aditivas para características de importancia económica como el contenido de azúcar, fibra y la resistencia a enfermedades; en tanto que la producción de biomasa, tiene similares valores de varianza genética aditiva y no aditiva, lo que es una excepción. En este trabajo se observó que, con respecto a este tema, las variables relacionadas con el azúcar son variables que parecen ser más útiles en la valoración de los genotipos **flor** que lo que podría ser en los genotipos **no_flor**. Lo anterior es más útil por cuanto es deseable incrementar los contenidos de sacarosa al inicio de los períodos de cosecha, ya que los genotipos florecedores se emplean para los inicios de cosecha y se desea que tengan altos valores de acumulación de sacarosa (Jackson, 2005).

5.3 HEREDABILIDAD

Los valores de heredabilidad se emplearon como factores de ponderación para el cálculo de los índices de selección propuestos por Smith (1936), Kempthorne y Nordskog (1959). En estas propuestas de “índices de desempeño” se emplea usualmente ponderadores de orden económico; sin embargo, debido a las complicaciones que requiere hacer un exhaustivo

análisis para establecer la importancia económica de todas las 14 variables en estudio, se prefirió utilizar los valores de heredabilidad, los cuales son aceptables para este caso (Williams, 1962; Nia *et al.*, 2011; Pérez-Elizalde *et al.*, 2014; Almeida *et al.*, 2014; Barbosa *et al.*, 2014). Así mismo, los valores de heredabilidad podrían ser buenos ponderadores ya que son derivados directamente del material vegetal en evaluación, solo variarán en función del material mismo y no dependerán de otros factores externos como el mercado. Si existe una desventaja, es que los valores de heredabilidad deben ser calculados para cada conjunto de datos, lo cual no es tan complicado si se compara con lo que se necesita para calcular ponderadores económicos (Smith 1936; Jackson *et al.*, 2015). Por otro lado, al parecer el nivel de uso de los índices Smith, KN, ESIM, y RESIM pueden estar relacionados con dos temas:

- a. Los valores de heredabilidad (Tabla 4.7) pueden ponderar con más peso a algunas variables que pueden no ser de gran valor en la discriminación de genotipos. Esto puede ser resuelto parcialmente, al identificar aquellas variables que no aportan considerablemente a la función y, en consecuencia, mejorar la calidad de estos índices a partir de evaluaciones de los índices mediante el uso de un *Path Analysis* (Sánchez, 2013; Ortiz *et al.*, 2015; Viveros *et al.*, 2015).
- b. Estos índices parecen funcionar mejor si las variables de las cuales se derivan guardan un importante nivel de correlación, como los componentes de rendimiento que se utilizan en granos (Bernardo, 1996; Ortiz *et al.*, 2015; Nia *et al.*, 2011). En el presente caso, los niveles de correlación entre las 14 variables utilizadas en la mayoría de los casos son bajos (Tabla 4.5).

5.4 ANÁLISIS DE COMPONENTES PRINCIPALES (ACP) GLOBAL

Las diferencias en términos de expresión fenotípica entre los conjuntos de datos **flor** y **no_flor** resultaron determinantes en la agrupación de las variables según los resultados del ACP, considerando todos los datos de las 14 variables incluidas (Tabla 4.7 y las Figuras 4.13 y 4.14). En este sentido, se sigue que, una buena expresión fenotípica en caña de azúcar, está definida por las condiciones de cultivo, es decir que la interacción *genotipo* \times *ambiente* ($G \times E$), la cual es determinante para obtener los mejores resultados en términos de mejora genética. Por tanto, es preferible, principalmente en los estados tempranos de selección, evaluar los genotipos en buenas condiciones de crecimiento a fin de

identificar apropiadamente aquellos genotipos que puedan tener potencial de convertirse en una variedad de uso comercial. De otra forma, éstos se perderían irremediablemente debido a la alta dinámica que se imprime en el proceso de mejoramiento genético en el Programa de Variedades de CENGICANÑA. Se ha demostrado que bajos niveles de aporte genético, unido a altos valores de heredabilidad, están asociados más bien a los efectos no aditivos, es decir al ambiente (Nagarajan *et al.*, 2006; Aitken *et al.*, 2008), por lo que un ambiente desfavorable no permitirá una evaluación apropiada. Por otro lado, la estabilidad agronómica de los cultivares es complicada de establecer y para esto se requiere de estudios a mayor escala sobre la interacción $G \times E$.

El vector correspondiente a la variable Población por metro lineal (**pob_9**) en la CP1, tuvo dirección contraria a los vectores de las variables relacionadas con la concentración de azúcar, lo que estuvo en concordancia con la correlación entre las altas poblaciones y las bajas concentraciones de azúcar (Figura 4.13), tema que se discutió con más amplitud en la sección de correlaciones (epígrafe 5.2).

5.4.1 Consideraciones generales sobre el ACP para todos los conjuntos de datos

En general para todos los datos de todos los conjuntos y en particular para cada uno de los 7 conjuntos de datos, se observó un patrón que indica que las primeras cuatro CPs contienen a la mayoría de las variables de importancia y, en consecuencia, según sea el caso, pueden ser empleadas como índices de desempeño para las variables correspondientes. Se estableció que entre el 46% y el 53% de la variabilidad de los conjuntos de datos analizados es explicada por las tres primeras CPs. Estas permiten hacer derivaciones apropiadas del total de la variabilidad que presentaron los siete conjuntos de datos analizados. Por otro lado, en los conjuntos de datos de genotipos “no florecedores” puede prescindirse, para los análisis respectivos, de la variable **flor** (%) puesto que no agrega ninguna variabilidad entre los genotipos en evaluación, ya que casi todos los genotipos tienen tasas bajas o nulas de floración, y en consecuencia, no hay variabilidad. Tena *et al.* (2016a) encontraron hasta el 61% de variación acumulada en las tres primeras CPs en el estudio de hasta 21 variables. Por otro lado, otros estudios han variado en diferentes combinaciones las variables que se utilizan para la formación de índices o ACP, y de esta forma han identificado hasta un 80% de explicación en las tres primeras CPs (Ong’ala *et al.*, 2016), o bien han desarrollado más de un

índice de selección al incluir o excluir variables en los modelos correspondientes (Singh y Khan, 1998; Priya *et al.*, 2013).

Las variables **oqueleve_3** y **pop_9** no están claramente contenidas en ninguna de las CPs y suelen estar mezcladas con variables que no son de su misma naturaleza; adicionalmente, los valores de estos dos vectores son bajos para la mayoría de funciones. Por tanto, es posible considerar que su participación genera cierta desestabilidad, por lo que pueden no ser útiles en el proceso de desarrollo de índices de selección.

Si bien se presentan diferencias entre los conjuntos de datos, se puede inferir que consistentemente la primera CP está asociada a características de acumulación de sacarosa, y que las CP2 y CP3 suelen estar más asociadas a las variables de acumulación de biomasa o bien con variables sobre la calidad interna de los tallos (oquedad y corcho), lo cual se presentó en forma alternada. Ong'ala *et al.* (2016) encontraron que las variaciones entre observaciones fueron de naturaleza diferente y no responden a un patrón similar al que se presentó en el presente trabajo, si bien la variable °Brix se asoció a la CP1. En todo caso, muchas de las variables medidas en el estudio de Ong'ala *et al.* (2016) fueron de naturaleza diferente.

Tena *et al.* (2016a) encontraron que para la CP1 las variables asociadas fueron relacionadas con la población en la unidad experimental. Esto puede estar relacionado con que en el trabajo de Tena *et al.* (2016a) se midieron muchas variables relacionadas con este tema en el modo de variables longitudinales y, en consecuencia, disminuyen las contribuciones de cada CP a la varianza total. Por otro lado, Ftwi *et al.* (2016) sí encontraron un patrón similar a los que se obtuvieron en este trabajo. En el estudio de Ftwi *et al.* (2016) el número de variables fue menor (12) y no contuvo variables longitudinales.

5.4.2 Análisis de modelos lineales mixtos para CPs para cada conjunto de datos

Jolliffe (2002) indica que las CPs son definidas como una transformación “ortonormal” (*orthonormal*) de \mathbf{x} , el cual es el vector-columna de las p variables aleatorias incluidas en el ACP. En este sentido, cada CP es una variable que puede ser sujeta de otros

tipos de análisis como el de conglomerados (*cluster analysis*), regresión mediante el uso de cuadrados mínimos (Massy, 1965), y otros análisis de orden paramétrico. Con esta base se discuten los resultados siguientes.

En el conjunto de datos **litoral_no_flor_plantía** fue donde se encontraron mayores diferencias significativas (Tabla 4.17). Esto puede deberse a que las condiciones de temperatura, humedad, radiación global y calidad del suelo en el estrato litoral propiciaron mejor desarrollo vegetativo (Watt *et al.*, 2014) y, en consecuencia, permitieron establecer con mayor precisión las diferencias de expresión de los genotipos en evaluación, según ya se mencionó, lo que es concordante con lo que se ha registrado en los campos comerciales de producción (CENGICAN 2011-2017).

Para el conjunto de datos **medio_flor_plantía** se presentaron algunas diferencias donde la significación no fue tan alta (Tabla 4.17), debido principalmente al escaso desarrollo fenotípico, como ya se apuntó. En todo caso, se cuenta con suficiente información respecto de las CPs, índices de desempeño y el modelo mixto, los cuales en su conjunto contribuyen con la diferenciación entre variedades y proporcionan información que contribuye a un proceso de selección un poco más preciso; si bien para este conjunto de datos, la mejor opción puede ser la selección por competencia, ya que ningún índice ponderó apropiadamente el valor genético de los materiales en estudio (Tabla 4.32).

5.4.3 Análisis Combinado de Varianza para los CPs

En el análisis combinado de varianza de todos los datos (Tabla 4.17), entre las CP1, CP2, CP3 y CP4 se observó que la CP1, en todos los casos, presentó buenas características de diferenciación, por lo que puede utilizarse como un índice de desempeño. Sin embargo se encontró que la dificultad de que, si las “cargas” (*eigenvectors*) de las principales variables relacionadas eran negativas, era preciso considerar que se invierta el índice de desempeño: los valores más bajos reflejan las mejores condiciones para cada observación. Con cargas positivas, ocurre lo contrario.

No se encontraron trabajos similares sobre el uso de los *scores* del ACP como índices de desempeño o de selección. Cerón-Rojas *et al.*, (2006 y 2008) sugieren el empleo de los coeficientes de las CPs como ponderadores para la generación de índices de desempeño,

los cuales se tratarán más adelante en la sección correspondiente. El uso más generalizado en un post-analysis de un ACP es la regresión de CPs (Massy, 1965; Fujiwara *et al.*, 2013).

5.5 MANOVA

La información de la Tabla 4.21 indica que, en conjunto, todas las variables de respuestas como un vector de medias diferencian significativamente a los genotipos. El principal aporte de este análisis es que se puede analizar los genotipos de caña desde un punto de vista multidimensional, y las diferencias entre variedades de caña de azúcar se pueden perfilar en términos también multidimensionales. En todo caso, el aporte de otras técnicas de análisis multidimensionales (ACPs por ejemplo), pueden complementar el MANOVA y ayudar a la selección de las variables más influyentes en cada conjunto de datos analizados.

Ong'ala *et al.* (2016) también encontraron, para diferentes combinaciones de variables (modelos), una elevada respuesta entre genotipos y que cuantos menos caracteres se incluyeron en el vector de medias de variables de respuesta, mayores fueron las diferencias entre clones. Ftwi *et al.* (2016) encontraron similares resultados en varios tipos de combinaciones de variables para diferentes MANOVAs, con los estadísticos de Pillai y Wilks.

5.6 ÍNDICES DE SELECCIÓN

Después de realizar, en los epígrafes precedentes, consideraciones generales, y tras establecer que se detectaron diferencias genotípicas entre genotipos para todos los conjuntos de datos en prácticamente todos los análisis, tanto unidimensionales como multidimensionales, se procede a discutir los **Índices de Desempeño** o **Índices de Selección**. La razón principal para llamarlos de **Desempeño** obedece a que, después de las experiencias obtenidas y los datos recabados y analizados, los índices estudiados funcionan más bien como un auxiliar en la toma de decisiones sobre la selección o no, de un genotipo en particular. En este sentido, los índices en cuestión no son indicadores categóricos para realizar o no, una selección. Por lo anterior no podrían considerarse, en un sentido estricto, índices de selección.

5.6.1 Correlaciones entre Índices

Para el conjunto de datos **medio_flor_plantía** se ha encontrado una elevada correlación entre TAH y BLUP, algo lógico ya que BLUP es el estimador insesgado del valor genético de los factores en estudio con base en el primero (Tabla 4.22) (Henderson *et al.*, 1959; Robinson, 1991). Así mismo, se encontraron correlaciones del TAH y BLUP con las CP2 y CP3, ya que éstas están asociadas a las variables relacionadas con la acumulación de biomasa, la cual es una parte importante para el cálculo del índice TAH (ver en el capítulo 3, Materiales y Métodos).

Por otro lado, el alto nivel de correlación entre ESIM y RESIM ($r = 1$) indica que esencialmente representan lo mismo y, por tanto, son intercambiables. Comparaciones similares se han realizado por medio de coeficientes de coincidencia (De Azeredo *et al.*, 2017; Ali y El-Sadek, 2016), los cuales estiman las frecuencias coincidentes de selección por parte de las diferentes metodologías.

Para los conjuntos de datos **medio_no_flor_1ª_soca** y **litoral_no_flor_plantía**, las tendencias de correlación entre índices fueron similares a las discutidas en este párrafo para **medio_flor_plantía**.

5.6.2 Índices Smith y Kempthorne y Nordskog

Con respecto a los resultados para el conjunto de datos **medio_flor_plantía**, se encontró que la varianza del índice en KN es más baja que la de Smith. En este sentido, es difícil establecer cuál de los dos índices es el de mejor calidad. En todo caso KN presenta una menor correlación con respecto al VMG, lo cual podría ser un indicativo de que Smith, con una mayor cercanía con dicho valor ($r=0.999$), es el que explica mejor los componentes genéticos y, en consecuencia, sería un mejor estimador del valor genético de los materiales en evaluación.

Para el presente caso, si bien hay buen nivel de heredabilidad en varias variables, la selección por KN es errática y pondera mejor a los genotipos con los valores medios y bajos de concentración de azúcar (**brixprom_1**, **pol_6**, **brixjugo_7** y **purity_8**) y valores bajos de peso (**peso5stalk_2**), por lo que no ofrece un ajuste apropiado con los objetivos de fitomejoramiento.

Algo similar ocurre con Smith. En varios otros estudios, Smith parece funcionar bien cuando se buscan y se encuentran algunas variables que puedan correlacionar bien entre sí y entrega un buen valor de ganancia genética, comparada con la selección directa de una sola variable, el TCH por ejemplo (Singh y Khan, 1998; Tahir *et al.*, 2014). Casos similares ocurren con otro tipo de índices (Priya *et al.*, 2013).

Por otro lado, es importante considerar que el empleo de los valores de heredabilidad como ponderadores para Smith y KN puede ser una complicación si, para un conjunto de datos en particular, las variables que más influyen (mayor valor de heredabilidad) en el modelo no son las que se emplean tradicionalmente para la valoración de los materiales genéticos (Ver Materiales y Métodos, “base de comparación”, epígrafe 3.3.6.7). En este sentido, la aplicación de un análisis de senderos (*path analysis*) puede ayudar a establecer qué variables son menos informativas para no ser utilizadas (Sánchez, 2013; Viveros *et al.*, 2015).

Almeida *et al.* (2014) encontraron que el uso de la heredabilidad como ponderador económico fue similar al uso de otros ponderadores como la desviación estándar genotípica o el coeficiente de variabilidad genotípica; sin embargo, las ganancias genéticas fueron calculadas sobre la base de un solo carácter y no de todos en su conjunto.

Jackson (2005), Hu *et al.*, (2008) y Jackson *et al.* (2015) sugirieron el uso de ponderadores económicos, los cuales pueden extraer con más amplitud las diferencias de las ganancias genéticas.

Los valores de heredabilidad pueden ser un buen sustituto del ponderador económico si la cantidad de caracteres en uso se reducen a aquellos que sean de verdadera importancia en la calificación fenotípica de los materiales genéticos en estudio (Almeida *et al.*, 2014; De Azeredo *et al.*, 2017, Viveros *et al.*, 2015).

Para el conjunto de datos **medio_flor_plantía**, se encontró que una baja expresión fenotípica en términos de las variables de crecimiento (biomasa) no permite una adecuada diferenciación entre genotipos. Lo anterior propició que las técnicas estadísticas no logaran establecer apropiadamente el valor genético de los materiales en evaluación y, en consecuencia, la diferenciación de las mismas fue, en general, poco consistente a través de los resultados obtenidos y discutidos para este conjunto de datos.

Si se estima que se ha alcanzado entre el 53% y 69% del potencial del rendimiento de biomasa de la caña de azúcar (Inman-Bamber, 2014), de acuerdo a la producción comercial de caña de azúcar (CENGICANA, 2011-2017), el potencial promedio mínimo para Guatemala debiera ser de 160 t ha⁻¹. En este sentido, con rendimientos muy bajos, la expresión real de la caña no se alcanzó en este conjunto de datos y, en consecuencia, no hubo diferenciación.

En estudios recientes (Tahir *et al.*, 2014), aplicaron el índice de Smith mediante un *path analysis*, con el objeto de aumentar el número de variables a incluir en el modelo. En general en este trabajo, la ponderación de variables de aumento de la biomasa son los que más influyeron en el valor del índice.

De Azeredo *et al.* (2017) encontraron que las selecciones directas por variable cambian de tal forma que no es posible hacer selecciones coincidentes, y que Smith no logra identificar apropiadamente buenos genotipos.

Otros índices funcionaron bien en la identificación de genotipos de superior desempeño, al utilizar los valores de heredabilidad como ponderadores económicos, según la metodología de Pesek y Baker (1970).

En el conjunto de datos **medio_no_flor_1ª_soca**, los índices SMITH y KN sugirieron selecciones con niveles bajos de desempeño. Esto puede ocurrir quizá por la debilidad del procedimiento en identificar correctamente los coeficientes apropiados para las funciones desarrolladas. En todo caso no se ajustaron apropiadamente con lo que se requiere en términos de mejora genética (la competencia del fitomejorador).

Análogamente, en el conjunto de datos **litoral_no_flor_plantía**, los índices Smith y KN no identificaron apropiadamente los genotipos con los mejores desempeños. Estas funciones pueden estar siendo influenciadas por los valores de heredabilidad, usados como sustitutos de ponderaciones económicas y, por tanto, asignaron un valor genético preciso, pero con ponderaciones en variables de importancia menor. En consecuencia, las selecciones sugeridas no fueron las que corresponden con genotipos de alta calidad. En este sentido, se insiste una vez más que un estudio de senderos (Path Analysis) podría ayudar a discernir sobre las variables que mejor califiquen el valor genético de los materiales en estudio (Sánchez, 2013).

5.6.3 ESIM y RESIM

Cerón-Rojas *et al.*, (2008b) indicaron que **RESIM** tiene la ventaja sobre **ESIM** de no requerir ponderaciones económicas (en este caso heredabilidades) ni de las estimaciones de la matriz varianzas-covarianzas. A pesar de algunas de las diferencias en las bases teóricas sobre las cuales se calculan estos índices (Cerón-Rojas *et al.*, 2006; Cerón-Rojas *et al.*, 2008b), en todos los conjuntos de datos, estos dos índices resultaron prácticamente intercambiables, es decir, presentaron prácticamente los mismos resultados de selección. Esto puede explicarse desde el punto de vista de que las variables de restricción en RESIM, las cuales pudieron ser las que más influyeron en las determinaciones de ESIM.

Los índices **ESIM** y **RESIM**, para el conjunto de datos **medio_flor_plantía**, tal como se muestra en la Tabla 4.32, guardan poca relación con la base de comparación de los índices: la *Selección por Competencia*. El problema en este conjunto de datos es que ningún índice guarda una asociación importante con la *selección por competencia*.

Otros autores prefieren hacer la comparación de eficiencia del índice con una estrategia de selección más puntual, como la de un solo carácter, el TCH o TAH por ejemplo (Pillai y Ethirajan, 1993; Miller *et al.*, 1978; Singh y Khan, 1998) o la selección directa por algún otro carácter en particular (De Azeredo *et al.*, 2017). La selección por competencia es más compleja e incluye muchos otros aspectos por lo que, alcanzar una alta eficiencia resultó difícil para todos los índices de selección. No hay referencias sobre la comparación con una metodología similar.

También en los datos de los conjuntos **medio_no_flor_1ª_soca** y **litoral_no_flor_plantía**, se encontró que los índices **ESIM** y **RESIM** fueron intercambiables. En general, las selecciones propuestas por estos índices ponderaron altamente la concentración de azúcar y, en términos intermedios, la acumulación de biomasa y la calidad del tallo. De lo anterior se sigue que las selecciones que son propuestas por estos dos índices son de materiales de un desempeño intermedio y no se ajustan a los requerimientos específicos de las estrategias de mejoramiento.

No se encontraron referencias sobre la utilización de un ACP como herramienta principal para la estructuración de un índice de desempeño en plantas o animales como lo proponen Cerón-Rojas *et al.* (2006) y Cerón-Rojas *et al.* (2008b), sino más bien como herramientas de

caracterización de las variables o valores a emplear como instrumentos de apoyo a la selección (Ali y El-Sadek, 2016) o para estudios de estabilidad mediante gráficos bi-plot (De Oliveira *et al.*, 2014; Quemé *et al.*, 2007, 2010).

5.6.4 CPs Como Índices de Desempeño

Con respecto a la selección mediante el uso directo de las CPs, el orden de prioridad de selección, para los siete conjuntos de datos evaluados, indica que primero deben seleccionarse los mejores genotipos según su concentración de azúcar y, dentro de éstos, a los genotipos con el mejor desempeño en cuanto a la acumulación de biomasa y/o calidad interna de los tallos (CP2 y CP3).

En un estudio de senderos (*path analysis*) Viveros *et al.* (2015) indicaron en sus hallazgos, de forma análoga, que la primera variable de estudio debería ser el contenido de azúcar, seguido por variables relacionadas con la biomasa (diámetro y largo del tallo). Si bien puede invertirse el orden de selección, se corre el riesgo de que no se cuente con genotipos con un buen desempeño en cuanto a la acumulación de sacarosa, lo cual es fundamental.

El mejoramiento primario por contenido de azúcar es prioritario a fin de equilibrar los gastos marginales asociados al corte, alce, transporte y molienda, según se indicó arriba (Jackson, 2005; Inman-Bamber *et al.*, 2011; Berding *et al.*, 2004). Las ganancias genéticas en producción global de azúcar (TAH), como producto del fitomejoramiento, se han dado por la vía de la acumulación de la biomasa, si bien estas ganancias año a año son difíciles de establecer como resultado de no poder diferenciar entre las ganancias por manejo agronómico, mejora genética o por condiciones climáticas favorables (Inman-Bamber, 2014); por otro lado, las ganancias relacionadas con respecto al contenido de azúcar se han mantenido estáticas (Inman-Bamber, 2014; Jackson, 2005). En este sentido, se hace muy importante privilegiar el estudio y selección en primer lugar por el contenido de sacarosa. La concentración de azúcar es uno de los más importantes caracteres en mejoramiento genético de caña de azúcar (Ming *et al.*, 2006; Berding *et al.*, 2004; Orozco *et al.*, 2012).

En el conjunto de datos **medio_no_flor_1ª_soca**, las selecciones mediante el uso de las CPs fueron una opción al procurar el equilibrio de selección entre genotipos con un buen contenido de azúcar y luego, dentro de estas selecciones, aquellas con el mejor desempeño en términos de acumulación de biomasa; sin embargo, las sugerencias de selección se ajustaron

poco con los intereses de mejoramiento (Tabla 4.45). Por otro lado, para el conjunto de datos **medio_flor_plantía** se encontró que las CPs asignaron un nivel aceptable de calificación a los genotipos; sin embargo, también se encontró muy poca asociación con la base comparadora de selección (selección por competencia) (Tabla 4.32). A pesar de lo anterior, las CPs plantearon un buen nivel de jerarquización, de tal manera que ésta pudo hacerse anidada según los *scores* promedio para cada variedad.

Análogamente, en el conjunto de datos **litoral_no_flor_plantía**, los índices basados directamente en las CPs lograron identificar, de forma equilibrada los genotipos con una buena expresión en concentración de azúcar, y después con respecto a la biomasa. Sin embargo, este fenómeno no identificó a los genotipos que en su totalidad entregaban más azúcar que el resto como consecuencia de una mayor acumulación de biomasa.

Como se indicó en el epígrafe 5.4.1, estudios similares en ACP agruparon en el primer componente principal a variables diversas y, en ocasiones, sin conexión aparente (Ong'ala *et al.*, 2016); sin embargo, otros estudios plantearon un agrupamiento similar a los que se presentaron en este estudio (Ftwi *et al.*, 2016). En este estudio, en todos los casos, las variables asociadas a la concentración de azúcar explicaron mayoritariamente la primera componente principal.

5.6.5 TAH y BLUP

Con respecto al índice TAH en el conjunto de datos **medio_flor_plantía**, se encontraron valores que indicaron que el ambiente tuvo una elevada influencia en esta variable empleada como índice, puesto que estuvo relacionado con variables de orden cuantitativo con una elevada interacción con el ambiente de evaluación. Además, es posible que estos resultados fueran también el reflejo del bajo nivel de expresión fenotípica de todos los materiales en evaluación.

Se sabe que las mediciones de los potenciales productivos de la caña están alrededor de las 100 t ha⁻¹ (Monteiro y Sentelhas, 2014; Ming *et al.*, 2006; Berding *et al.*, 2004; Inman-Bamber, 2014). Inman-Bamber (2014) indicó que los potenciales en TAH pueden rondar en el trópico entre las 33 y 35 t ha⁻¹; sin embargo, alcanzar este potencial está más asociado a la biomasa. Si bien también Inman-Bamber (2014) indicó que el potencial de la caña de azúcar

en biomasa está entre el 53% y el 69 en los actuales momentos, producir variedades de caña de azúcar con altos potenciales de rendimientos es cada vez más complicado.

Jackson (2005) indica que existe la posibilidad de haber ya identificado y explotado los alelos relacionados con el contenido de sacarosa, de ahí el estancamiento en el crecimiento con respecto a este carácter, lo cual también es discutido por Inman-Bamber *et al.* (2011); es decir, es posible que ya se haya alcanzado el potencial productivo de la caña de azúcar, al menos en este carácter.

Al considerar al TAH como un índice de desempeño, y de acuerdo con los resultados de un ANOVA y con la comparación de estos resultados con el BLUP presentados en la Tabla 4.31, se encontró que para efectuar la selección es razonable decantarse por el empleo de BLUP, como la herramienta que ayude a discriminar mejor a los genotipos y, en consecuencia, auxilie a realizar una selección más precisa según se acotó en los apartados anteriores en este epígrafe. El uso de BLUP está siendo cada vez más utilizado en la actualidad en mejoramiento genético de plantas (Piepho *et al.*, 2008).

En caña de azúcar se ha propuesto su uso derivado también de una variable similar, Toneladas °Brix ha⁻¹ (De Oliveira *et al.*, 2013) y para TCH (Barbosa *et al.*, 2014). Los valores de BLUP que se discuten aquí fueron derivados de la variable TAH. BLUP es un estimador de los efectos aleatorios en un modelo lineal mixto (Piepho *et al.*, 2008), en el cual todos los factores, incluido el pertinente a genotipos, se consideran aleatorios (Thompson, 2008; Robinson, 1991; Piepho *et al.*, 2008; Bates *et al.*, 2015).

En el conjunto de datos **medio_no_flor_1a_soca**, los índices que más se aproximaron a los valores esperados en un sistema de selección por competencia fueron TAH y BLUP. Esto, como ya antes se apuntó, es relativamente normal ya que el TAH es una de las más importantes variables para hacer las selecciones por competencia; sin embargo es preciso anotar que en todo caso la asociación no presentó un elevado nivel de correlación (0.34 y 0.33 para TAH y BLUP, respectivamente, Tabla 4.45). Para el caso de este conjunto de datos, las expresiones fenotípicas fueron de mayores niveles (Tablas 4.1 y 4.2). Este fenómeno permitió que las diferenciaciones entre genotipos fueran más clara y consistentes según las repeticiones. Fue posible identificar genotipos con un alto potencial productivo y éstos superaron claramente a los testigos establecidos dentro del ensayo (Tabla 4.43).

Adicionalmente, en este grupo de datos, según la regresión **Logit**, los BLUPs presentaron los mejores niveles de asociación con los requerimientos de mejora genética, o de requisitos a cumplir según la competencia de los fitomejoradores, si bien estos valores fueron intermedios.

En los datos del conjunto **litoral_no_flor_plantía**, el TAH identificó con precisión a los genotipos sobresalientes como consecuencia de la combinación de los valores de acumulación de biomasa y concentración de azúcar. En este sentido se presentó como uno de los mejores indicadores del valor genético de los genotipos en evaluación. Por otro lado, en este mismo conjunto de datos, BLUP permitió una corrección de las tendencias de calificación de los genotipos y, en consecuencia, corrigió la sugerencia de selección, de tal forma que se pudo ser más preciso en el proceso de valoración genética (Henderson *et al.*, 1959, Robinson, 1991; Thompson, 2008).

La variación en la selección, ya sea por TAH o el BLUP, puede generar selecciones diferenciales. Esto es particularmente importante, puesto que la variable TAH puede presentar algunos problemas en términos de la correcta estimación del valor genético de los materiales en estudio. La variación en las lecturas del TAH, según los valores medidos en las repeticiones y localidades, puede presentar una variación importante, y ésta puede ser corregida a través del cálculo de BLUP.

Las estimaciones de BLUP permitieron corregir algunas desviaciones en TAH y permitieron presentar un buen nivel de selección después de las rectificaciones planteadas. El cálculo de BLUP sobre TAH presentó los resultados más consistentes de selección, si bien su asociación con la base de comparación fue de niveles intermedios. Su calidad y precisión no presentó ventajas notables con respecto a TAH.

Si bien el uso de BLUP es rutinario en mejoramiento genético de animales, las referencias del uso de BLUP es limitada en la mejora genética de plantas, y no es sino hasta inicios del siglo XXI que se ha empezado a hacer un uso más intensivo de este instrumento estadístico en la mejora genética de plantas (Bernardo, 1996; Simeão *et al.*, 2002; Bauer *et al.*, 2006; Carvalho *et al.*, 2008; Piepho *et al.*, 2008; Crossa *et al.*, 2010; Chiorato *et al.*, 2008), principalmente para la predicción de los valores genéticos (*breeding values*) y la estimación de parámetros genéticos. En caña de azúcar las aplicaciones al respecto son más limitadas aún (Barbosa *et al.*, 2014; De Oliveira *et al.*, 2013), y se ha empleado principalmente en el estudio

del desempeño de progenies con respecto a sus progenitores en estados de evaluación iniciales.

En el presente estudio, BLUP asociado al TAH se contempló también como un índice de desempeño, si bien en realidad es el estimador del valor genético de los genotipos en evaluación (Henderson *et al.*, 1959; Robinson, 1991; Thompson, 2008). En este estudio, mediante el uso de BLUPs, de todas las observaciones de los siete conjuntos de datos (Anexo 3), se identificaron en la fracción superior (10%) 25 genotipos cuyo promedio de TAH fue de 28.10 t, que equivale al 22% de ganancia genética con respecto a los dos principales testigos empleados: CP72-2086 y CP73-1547.

5.6.6 Regresiones Logit

En el conjunto de datos del estrato **medio_flor_plantía**, las regresiones fueron bajas, debido principalmente al bajo desempeño fenotípico de todos los materiales genéticos, según se discutió anteriormente, lo que estuvo relacionado con que no se alcanzó un elevado valor de acumulación de biomasa según el potencial (Monteiro y Sentelhas, 2014; Inman-Bamber, 2014), lo que derivó en bajos niveles de expresión y, en consecuencia, en un bajo nivel de diferenciación fenotípica entre genotipos. Debido a lo anterior, se sigue que ningún índice ofreció una buena opción para el auxilio en la selección. Una regresión **Logit** (Tablas 4.32, 4.45 y 4.56) donde se cotejaron los índices con la base comparación (competencia del fitomejorador), los valores de correlación con los índices Smith y KN fueron prácticamente cero, lo que es muestra de que, tanto Smith como KN, tuvieron muy poca relación con el interés de selección del fitomejorador, lo cual fue similar para el resto de índices.

En el conjunto de datos **medio_no_flor_1ª_soca**, los índices **ESIM** y **RESIM** parecieron asociarse apropiadamente con el VMG, las selecciones propuestas por estos índices no fueron las mejores opciones, lo cual estuvo relacionado con la asociación entre los índices en cuestión y la base de comparación (regresión **Logit**). Por lo anterior, es razonable la baja asociación entre **ESIM** y **RESIM** con TAH principalmente, puesto que éste último sí considera importantes tanto a la acumulación de biomasa como a la concentración de azúcar.

En los conjuntos de datos del estrato **medio_no_flor_1ª_soca** y **litoral_no_flor_plantía**, los modelos con mejor correlación **Logit** con respecto a la base de comparación (selección por

competencia) fueron TAH y BLUP. La razón por la cual esta correlación elevada entregó valores elevados, es que la selección por competencia incluye al TAH como una variable de alto valor de ponderación en la toma final de decisiones según la competencia.

Sin embargo, para el caso del conjunto de datos **litoral_no_flor_plantía** todos los valores de correlación aumentaron considerablemente para varios índices, que en los casos precedentes presentaron niveles muy bajos. Una vez más, la explicación sobre este fenómeno puede abordarse desde la instancia de la expresión fenotípica. Una mayor expresión fenotípica, en casi todas las variables, está asociada a una diferenciación más puntual entre los genotipos y, en consecuencia, este fenómeno se refleja en las funciones que se generan para cada índice de desempeño. En consecuencia, también, es posible medir con más precisión y calidad el valor genotípico de los materiales. Esto se observó también en el tema de las heredabilidades para las variables en estudio (Tablas 5.7, 5.40 y 5.43), así como para las CPs y el TAH.

Otros estudios utilizaron la comparación de la eficiencia del índice contra la selección directa por alguna de las variables de estudio, contenido de fibra por ejemplo (Priya *et al.*, 2013; De Oliveira *et al.*, 2013; De Azeredo *et al.*, 2017); sin embargo, este no es el caso en mejora genética. El fitomejorador utiliza varios estándares de selección para varias variables y no solo para una. Es por eso que se decidió utilizar la variable de comparación propuesta en este estudio, la cual resume (según se indicó en la sección de Materiales y Métodos) la selección final de un genotipo de caña de azúcar. La desventaja aparente es que la variable de comparación es binomial por lo que la regresión se hizo en términos logísticos (Manning, 2007; Hu *et al.*, 2006). En todos los casos las eficiencias más altas (Tablas 4.32, 4.45, y 4.56) se adjudicaron a la valoración genotípica realizada por los BLUPs y el TAH, de lo que se sigue que fueron los valores que mejor revelaron los valores genéticos de los materiales de caña de azúcar en estudio.



6. CONCLUSIONES



6 CONCLUSIONES

- 6.1. Los valores promedio para la acumulación de biomasa (**peso5stalks_2** y **alt_10**) y la producción de sacarosa (TAH) fueron consistentemente más elevados en el estrato litoral; al contrario, los valores promedio de concentración de azúcar (**pol_6**, **brixprom_1** y **brixjugo_7**) fueron significativamente más elevados en el estrato medio que en el litoral. Por tanto, se concluye que la selección de cultivares por su valor de biomasa y concentración de sacarosa debe ser contrastada en distintos ambientes.
- 6.2. Las diferencias de desempeño de los genotipos en términos de TAH fueron altamente significativas e indican que la componente genética, en seis de los siete conjuntos de datos, fue el factor más determinante para la diferenciación. Por otra parte, el factor de interacción *genotipo* \times *ambiente* indica la necesidad de seguir estableciendo ensayos comparativos en diversos ambientes para la evaluación de su expresión fenotípica. Además, el pobre desarrollo en biomasa en dos de los ensayos no permitió discriminar entre genotipos en ninguno de los índices aplicados y, por tanto, la aplicación de modelos de selección (cálculos de heredabilidad y otros), en estas condiciones, no fueron útiles como herramientas de diferenciación genética.
- 6.3. Las variables con mayor potencial de selección en un proceso de búsqueda de ganancia genética, atendiendo a los elevados valores de heredabilidad, fueron las relacionadas con el contenido de azúcar (**brixprom_1**, **pol_6**, **brixjugo_7**) y, en menor medida, las variables **peso5stalks_2**, **diam_11**, y **pith_13**.

- 6.4. Las variedades no florecedoras tendieron a concentrar más azúcar (**brixprom_1**, **pol_6** y **brixjugo_7**) que las variedades florecedoras, por lo que resultaron más interesantes para la mejora de caña de azúcar.
- 6.5. Las elevadas correlaciones (****) que se encontraron entre las variables relacionadas con el contenido de sacarosa en el tallo, en concreto el Brix medido en el campo (**brixprom_1**) con las mediciones del laboratorio (**pol_6** y **brixjugo_7**), permiten su utilización como un elemento preliminar de selección.
- 6.6. La variabilidad explicada por las tres primeras componentes principales del ACP varió entre el 50% y 55% para cada conjunto de datos analizado, y fue suficiente para explicar las tendencias y estructura de organización de los datos, así como su uso potencial como índices de desempeño.
- 6.7. La CP1 del ACP, en todos los conjuntos de datos, se asoció con las variables relacionadas con el contenido de sacarosa (**brixprom_1**, **pol_6**, **brixjugo_7** y **purity_7**) con un rango de la variabilidad explicada entre el 19.8% y el 24.3%. La CP2 estuvo influida por las variables relacionadas con la acumulación de biomasa (**peso5stalks_2**, **diam_11**, **alt_10**, y **fiber_14**). La CP3 estuvo relacionada con la calidad interna de los tallos (principalmente **flower_12** y **pith_13** y, en menor grado, por **fiber_14** y **oqueinter_3**). Las variables asociadas a las CP2 y CP3 se presentaron en una u otra dependiendo del ensayo, y agruparon entre el 11 y el 16% de la variabilidad total cada una.
- 6.8. Según la organización encontrada en los ACP, los análisis de varianza y funciones lineales de los índices de desempeño, las variables que más incidieron en la variación detectada fueron, en orden de importancia: **brixprom_1**, **pol_6**, **brixjugo_7**, **purity_8**, **peso5stalks_2**, **diam_11**, **alt_10**, **fiber_14**, **flower_12**, **pith_13** y **oqueinter_4**. La variable **pop_9** presentó inconsistencia generalizada y se agrupó en varios de los ACP en la CP1 o en la CP2. Consiguientemente, las variables listadas se presentaron como las que consistentemente explican las variaciones genéticas y, en consecuencia, pueden ser empleadas en estudios de diferenciación entre genotipos. Por otra parte, las variables **oqueleve_3** y **oqueseve_5** no tuvieron una influencia importante en términos de

caracterización de los genotipos en estudio, ya que su variación fue muy escasa y aportaron poca información, por lo que puede prescindirse de ellas sin pérdida importante de información. Por último, la variable **flower_12** fue útil en la diferenciación genética entre materiales, solamente en los experimentos donde se incluyeron variedades florecedoras. En el caso de los experimentos de variedades no florecedoras, las varianzas correspondientes a esta variable son prácticamente cero y, por tanto, no son útiles para discriminar genotipos en este tipo de experimentos.

- 6.9. El uso de al menos las tres primeras CPs del ACP, en términos de su aplicación como elementos de discriminación (índices de desempeño) entre genotipos de caña de azúcar, permitieron realizar la selección de genotipos de forma anidada. En nuestro estudio, en primer lugar teniendo en cuenta las variables relacionadas con el contenido en azúcar (CP1), seguidas en segundo lugar por las de biomasa (CP2) y, finalmente, en tercer lugar, por las de calidad interna del tallo (CP3). Por tanto, después de seleccionar el 25% de los genotipos según la CP1, se seleccionaron aquellos que tuvieron una mejor respuesta a la CP2 y a la CP3, respectivamente. Sin embargo, las CPs, empleadas como índices de desempeño, sacarosa, resultaron en sugerencias de selección de genotipos de valor general intermedio. Si bien las selecciones obtenidas por las CPs presentaron un buen nivel de desempeño con respecto a la acumulación de sacarosa, estas selecciones no siempre se correspondieron con genotipos con buenas lecturas en términos de acumulación de biomasa. Por tanto, se concluye que el uso de estos índices fue limitado y su uso se recomienda a conjuntos de datos donde se establezcan claras diferencias entre genotipos para las principales variables de valoración, como el contenido de azúcar, por ejemplo.
- 6.10. En un análisis multidimensional de variación (MANOVA) se encontró que las diferencias entre medias de todas las variables en su conjunto (como un vector de medias) fueron debidas a la componente genética; y, por tanto, la formulación *a priori* de contrastes pueden ser una prueba que justifique la separación de los tratamientos (genotipos) o bien por otros medios aplicables como BLUP.
- 6.11. En general, los índices mostraron valores genéticos que no reflejaron los verdaderos valores de cada genotipo, presentando valores bajos de correlación con el interés de

selección del fitomejorador. Consecuentemente, su utilidad en las condiciones del estudio fue notoriamente reducida.

- 6.12. Los índices Smith y KN no lograron ajustar la propuesta de selección con el verdadero potencial de los genotipos. ESIM y RESIM propusieron, en general, las mismas selecciones en prácticamente todos los casos y, por tanto, resultaron intercambiables. Esto es debido a que las variables de no restricción en RESIM fueron las de mayor carga en ESIM. En todo caso, las selecciones sugeridas correspondieron a genotipos de desempeño intermedio. Las variables de mayor ponderación fueron las relacionadas con la concentración de azúcar. Consiguientemente, la utilidad de estos índices también fue de reducido valor. ESIM y RESIM pueden ser apropiados para un conjunto pequeño de variables, que principalmente estén ligadas a las componentes del rendimiento. En este sentido su aplicabilidad es limitada en caña de azúcar, donde se estudian y se ponderan con importancia relativa un mayor número de variables, las cuales no necesariamente están correlacionadas todas entre sí.
- 6.13. De acuerdo con las regresiones **Logit**, los índices de desempeño Smith, Kempthorne y Nordskog, ESIM, y RESIM presentaron ajustes bajos con respecto al patrón de comparación, por lo que, en condiciones similares de estudio, no pueden ser utilizados rutinariamente.
- 6.14. El ajuste de los índices de selección estudiados en este trabajo no tuvieron un patrón generalizado (Smith, KN, ESIM, RESIM, PCs). Por tanto, el ajuste de estos índices, y probablemente otros tipos de índices, deben ser priorizados según cada conjunto de datos. En este sentido es imperativo indicar que el patrón y ajuste de datos será el que establezcan las condiciones y el marco de trabajo para la búsqueda y ajuste del índice más apropiado. Incluso, con un mismo conjunto genotipos, en dos o más localidades, las variaciones por la interacción *genotipo* \times *ambiente* mostrarán variaciones en cuanto al ajuste y el tipo de índice a emplear.
- 6.15. La selección de genotipos de caña de azúcar fueron función de los valores genéticos según el TAH y BLUP. Para el resto de índices no se encontró ninguna relación funcional. BLUP fue capaz de optimizar los procedimientos de selección de genotipos

de caña de azúcar y puede ser aplicado en forma rutinaria en los procesos de evaluación y selección de genotipos en condiciones de cultivo en Guatemala. En conclusión, el procedimiento estadístico más apropiado para diferenciar genotipos de caña de azúcar es el Mejor Estimador Lineal Insesgado (BLUP) a partir de la variable TAH. Este mismo estimador del efecto aleatorio de cada genotipo, puede emplearse en otras variables, pero estas pueden no contener los criterios generales de selección que TAH sí contiene.

- 6.16. No fue posible establecer una función que, en su conjunto, ponderara apropiadamente las 14 variables en estudio. En este sentido, es recomendable realizar ajustes de modelos por grupos diferentes de variables, a partir de los hallazgos en los ACP y el resto de funciones y análisis, en términos de la estructura de organización de las mismas.
- 6.17. Un buen desarrollo fenotípico se refleja en un mejor nivel de diferenciación de los genotipos y en un mejor discernimiento del valor genético de los materiales en evaluación. En los casos en donde los experimentos se encontraron con deficiencias en el desarrollo fenotípico, todos los procedimientos empleados para la diferenciación entre genotipos no fueron competentes en discernir apropiadamente entre materiales.
- 6.18. El desajuste de varios de los índices con respecto al verdadero interés de selección puede ser corregido mediante la reducción de la dimensionalidad. En este estudio fue posible identificar las variables con menor carga en las diferentes funciones, que fueron más o menos las mismas en todos los casos, y su eliminación de los análisis puede conducir a un mejor nivel de ajuste de algunos índices, particularmente RESIM.
- 6.19. De acuerdo con la conclusión anterior, diferentes tipos de análisis pueden ser empleados para el mejor aprovechamiento de la información, particularmente el análisis de senderos (*path analysis*).
- 6.20. La utilización de la metodología de BLUP permitió la selección en la fracción superior (10%) de 25 genotipos cuyo promedio de TAH fue de 28.10 t, que equivale al 22% de ganancia genética con respecto a los dos principales testigos empleados: CP72-2086 y CP73-1547.





7. REFERENCIAS BIBLIOGRÁFICAS



7 REFERENCIAS BIBLIOGRÁFICAS

- ADLER, D.; MURDOCH, D. Y OTROS (2017). rgl: 3D Visualization Using OpenGL. R package version 0.98.1. <https://CRAN.R-project.org/package=rgl>
- AITKEN, K.; JACKSON, P.; MCINTYRE, L. (2005). A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theor Appl Genet* 110:789–801.
- AITKEN, K.; JACKSON, P.; MCINTYRE, C. (2006). Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar x *Saccharum officinarum* population. *Theor Appl Genet* 112:1306-1317.
- AITKEN, K.; HERMAN, S.; KARNO, K.; BONNETT, G.D.; MCINTYRE, C.; JACKSON, P. (2008). Genetic control of yield related stalk traits in sugarcane. *Theor Appl Genet* 117:1191–1203.
- ALI, M.; EL-SADEK, A. (2016). Evaluation Of Drought Tolerance Indices For Wheat *Triticum aestivum* L.) Under Irrigated And Rainfed Conditions. *Communications In Biometry And Crop Science*. 11(1): 77–89.
- AL-JANABI, S.; HONEYCUTT, R.; MCCLELLAND, M.; SOBRAL, B. A genetic map o *Saccharum spontaneum* L. “SES 208”. (1993). *Genetics* 134:1249-1260.
- ALLARD, R. (1980). Principios de la mejora genética de las plantas. Barcelona. Omega. 498 p.
- ALMEIDA, L.M.; VIANA, P.A.; TEIXEIRA, A.; BARROS, J. (2014). Breeding full-sib families of sugar cane using selection index. *Ciência Rural*, Santa Maria, 44(4): p 605-611.

- BABU, C.; KODALINGAM, K.; NATARAJAN, U.S.; SHANTHI, R.M.; THANGASAM, Y. (2006). Cytological studies on sugarcane intergenérico hybrids. Second National Plant Breeding Congress. India. Plant Breeding in Post Genomics Era. Proceedings 206-209.
- BANCO DE GUATEMALA. (2015). Guatemala en Cifras. Banco de Guatemala. Guatemala. 71 p.
- BARBOSA, M.H.; FERREIRA, A.; PEIXOTO, L.A.; RESENDE, M.D.; NASCIMENTO, M.; SILVA, F.F. (2014). SELECTION OF SUGAR CANE FAMILIES BY USING BLUP AND MULTI-DIVERSE ANALYSES FOR PLANTING IN THE BRAZILIAN SAVANNAH. GENET MOL RES. 12;13(1):1619-1626. DOI: 10.4238/2014.MARCH.12.14.
- BARBOSA, R.P.; ALCANTARA, F.; GRAVINA, L.M.; GRAVINA, G.A.; PORTELA, M.G.; BEZERRA, A.A. (2017). Early selection of sugarcane using path analysis. Genetics and Molecular Research 16 (1): gmr16019038
- BATES, D. (2013). Computational Methods for Mixed Models. Universidad de Wisconsin. Madison. 22 p. Consultado 20 nov 2016. Disponible en <https://cran.r-project.org/web/packages/lme4/vignettes/Theory.pdf>.
- BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. (2015). Fitting Linear Mixed Effects Models Using lme4. Journal of Statistical Software, 67(1):1-48. doi:10.18637/jss.v067.i01
- BAUER, A.M.; REETZ, T.C.; LÉON, J. (2006). Estimation of Breeding Values of Inbred Lines using Best Linear Unbiased Prediction (BLUP) and Genetic Similarities. Crop Sci 46:2685-2691. doi:10.2135/cropsci2006.01.0019.
- BEGNA, S.H.; HAMILTON, R.I., DWYER, L.M.; STEWART, D.W.; SMITH, DL. (1997). Effects of Population Density and Planting Pattern on the Yield and Yield Components of Leafy Reduced-Stature Maize in a Short-Season Area. J. Agronomy & Crop Science. 179:9-17.
- BERDING, N.; HOGARTH, M.; COX, M. (2004). Plant Improvement of Sugarcane. *In* Sugarcane. Blackwell Science. Oxford, Reino Unido. p. 20:53.
- BERNARDO, R. (1996). Best Linear Unbiased Prediction If Maize Single-Cross Performance. Crop Sci 36:50-56.
- BERNARDO, R. (2008). Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. Crop Sci 48: 1649-1664.

- BIRKETT, H.S. (1998). Core sampling recommended procedures. Audubon Sugar Institute, Louisiana State University Agricultural Center. 16 p.
- BISSESSUR, D.; RAMNAWAZ, C.; RAMDOYAL, K. (2010). Use of historical multilocation sugarcane variety trials data to identify relationships among environments in terms of genotype responses. *Proc. Int. Soc. Sugar Cane Technol.* 27:1–9.
- BOS, I.; CALIGARI, P. (2007). *Selection Methods in Plant Breeding*, 2nd edition. Springer. 471 p.
- BROWN, T. (1999). *Genomes*. Wiley and Sons Inc. New York, 472 p.
- BULMER, M. (1980). *The Mathematical Theory of Quantitative Genetics. Lectures in Biomathematics*, University of Oxford: Clarendon Press. 254 p.
- BUTTERFIELD M.K.; D'HONT, A. (2001). The sugarcane genome: A synthesis of current understanding, and lessons for breeding and biotechnology. *Proc S Afr Sug Technol Ass* 75: 1-5.
- CAMADRO, E.L.; MENDIBURU, A.O. (1988). Utilización de germoplasma en el mejoramiento de papa. *Revista Latinoamericana de la papa*. 1 (1): 35-43.
- CARDOZO, NP; SENTELHAS, PC. (2013). Climatic effects on sugarcane ripening under the influence of cultivars and crop age. *Scientia Agricola*. 70(6):449-456.
- CAREY, G. 1998. *Multivariate Analysis of Variance (MANOVA): I. Theory*. Consultado 30 nov 2016. Disponible en <http://ibgwww.colorado.edu/~carey/p7291dir/handouts/manova1.pdf>.
- CARVALHO, A.D.; NETO, R.F.; GERALDI, I.O. (2008). Estimation and prediction of parameters and breeding values in soybean using REML/BLUP and Least squares. *Crop Breeding and Applied Biotechnology*. 8:219-224.
- CENGICAÑA (CENTRO GUATEMALTECO DE INVESTIGACIÓN Y CAPACITACIÓN DE LA CAÑA DE AZÚCAR). (2011-2017). Base de datos de productividad de la agroindustria azucarera guatemalteca. Guatemala. Consultado el 18 de Mayo de 2017. Disponible con restricciones en <https://cengicana.org/usodeingenios>
- CERÓN-ROJAS, J.; SAHAGÚN-CASTELLANOS, J. (2005). Un índice de selección basado en componentes principales. A selection index based on principal components. *Agrociencia* 39:667-677.
- CERÓN-ROJAS, J.; CASTILLO-GONZÁLEZ, F.; SANTACRUZ-VARELA, A. (2006). A selection Index method based in eigenanalysis. *Crop Sci* 46:1711-1721.

- CERÓN-ROJAS, J.; CASTILLO-GONZÁLEZ, F.; SANTACRUZ-VARELA, A.; BENITEZ-RIQUELME, I.A.; CROSSA, J. (2008a). A molecular selection index method based in eigenanalysis. *Genetics*. 180:547-557.
- CERÓN-ROJAS, J.; CASTILLO-GONZÁLEZ, F.; SANTACRUZ-VARELA, A.; CROSSA, J. (2008b). A restricted selection index based on eigenanalysis. *J. Agric. Bio. Environ Statist* 13:440-457.
- CHIORATO, A.F.; CARBONELL, S.A.; DIAS, L.A.; RESENDE, M.D. (2008). Prediction of Genotypic Values and Estimation of Genetic Parameters in Common Bean. *Brazilian Archives of Biology and Technology*. 51(3):465-472.
- COLLEAU, J.J.; DUCROCQ, V.; BOICHARD, D.; LARROQUE, H. (SF). Approximate multitrait BLUP evaluation to combine functional traits information. Institut National de la Recherche Agronomique. Francia. 9 p. consultado el 10 junio 2017. Disponible en http://agtr.ilri.cgiar.org/documents/Library/docs/Interbull/bulletin23_files/Docs/colleau2.pdf.
- CROSSA, J.; DE LOS CAMPOS, G.; PÉREZ, P.; GIANOLA, D.; BURGUEÑO, J.; ARAUS, J.L.; MAKUMBI, D.; SINGH, R.P.; DREISIGACKER, S.; YAN, J.; ARIEF, V.; BANZIGER, M.; BRAUN, H.J. (2010). Prediction of Genetic Values of Quantitative Traits in Plant Breeding Using Pedigree and Molecular Markers. *Genetics* 186:713–724.
- CULLIS, B.; SMITH, A.; COOMBES, N. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural Biological and Environmental Statistics*. DOI: 10.1198/108571106X154443.
- CUNNINGHAM, E.P.; MOEN, R.A.; GJEDREM, T. (1970). Restriccion of selection indexes. *Biometrics*, 26:67-74.
- D'HONT, A.; ISON, D.; ALIX, K.; ROUX, C.; GLASZMANN, J.C. (1998). Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome*, 41:221-225.
- DA SILVA, J.; SORRELLS, M.; BURNQUIST, W.; TANKSLEY, D. (1993). RFLP linkage map and genome analysis of *Saccharum spontaneum*. *Genome* 36:782-791.
- DA SILVA, G.O.; DE CARVALHO, A.D.; SOUZA, Z.; PONIJALEKI, R.S.; PEREIRA, A. (2016). Desempenho genotípico de clones de batata via modelos mistos. *Revista Brasileira de Ciências Agrárias*. DOI:10.5039/agraria.v11i4a5390.
- DE AZEREDO, A.; BHERING, L.; BRASILEIRO, B.P.; CRUZ, C.D.; SILVEIRA, L.C.; OLIVEIRA, R.A.; BESPALHOK, J.C.; DAROS, E. (2017). [Comparison between](#)

- [different selection indices in energy cane breeding](https://doi.org/10.4238/gmr16019535). Genet Mol Res. 15;16(1). doi:10.4238/gmr16019535.
- DE OLIVEIRA, R.A.; DAROS, E.; RESENDE, M.D.; FILHO, J.C.; ZAMBON, J.L.; RUARO, L. (2013). Early selection in sugarcane family trials via BLUP and BLUPIS procedures. Acta Scientiarum. Doi: 10.4025/actasciagron.v35i4.16430.
- DE OLIVEIRA, R.L.; VON PINHO, R.E.; FERREIRA, D.F.; PIRES, L.P.; MELO, W.M. (2014). Selection Index in the Study of Adaptability and Stability in Maize. The ScientificWorldJournal. Article ID 360570. DOI <http://dx.doi.org/10.1155/2014/360570>.
- DEEPAYAN, S. (2008). Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5.
- DUNTEMAN, G.H. (1984). Introduction to multivariate analysis. Sage publications. California. 237 p.
- EDMÉ, S.; COMSTOCK, J.C.; MILLER, J.D.; TAI, P.Y. (2005). Determination of DNA content and genome size in sugarcane. Journal American Society Sugar Cane Technologists 25:1-16.
- ENGELKE, J. (2002). Sugarcane: Measuring Commercial Quality. Department of Agriculture. Government of West Australia. No. 23/2002. Agdex 712/25.
- EVERITT, B.S. (2014). Logistic Regression and Generalized Linear Models: Blood Screening, Women's Role in Society, Colonic Polyps, Driving and Back Pain, and Happiness in China. In A Handbook of Statistical Analyses Using R. 3rd Edition. Viñeta 21 p.
- FAO. 2015. Statistical Pocketbook (2015). World Food and Agriculture. FAO, Roma. 231 p.
- FAS. USDA. (2017). Sugar: World Markets and Trade. Foreign Agricultural Service/USDA. Estados Unidos.
- FEDERER, W.T.; RAGHAVARAO, D. (1975). On Augmented designs. Biometrics. 31:29-35.
- FEHR, W. (1987). Principles of Cultivar Development. Volume 1. Theory and Technique. Macmillan Publishing Company. New York. 536 p.
- FISHER, R. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics. (7):179-89.

- FOOD AND AGRICULTURAL ORGANIZATION OF THE UNITED NATIONS. (2009). FAO. (consultado 29 mayo 2017). Disponible en <http://faostat.fao.org/default.aspx>.
- FRANCO, C.; ASTUDILLO, J.; BONILLA, P.; LÓPEZ, J. (2013). Mapping Qtl Associated With Sucrose And Biomass Content In Colombian Sugarcane Varieties (*Saccharum Spp*). Proc Int Soc Sugar Cane Technol, Vol. 28
- FTWI, M.; MEKBIB, F.; ABRAHA, E. (2016). Multivariate analysis of sugar yield contributing traits in Sugarcane (*Saccharum officinarum* L), in Ethiopia. African Journal of Plant Science. 10(8):145-156.
- FUJIWARA, M.; MINAMIDANI, T.; NAGAI, I.; WAKAKI, H. (2013). PRINCIPAL COMPONENTS REGRESSION BY USING GENERALIZED PRINCIPAL COMPONENTS ANALYSIS. JOURNAL OF THE JAPAN STATISTICAL SOCIETY. 43 (1): 57-78.
- FURLANI, R.; DE MORAES, M.; RESENDE, M.; JUNIO, E.; GONÇALVES, P.; FILHO, W. (2005). Estimation of variance components and prediction of breeding values in rubber tree breeding using the REML/BLUP procedure. Genetics and Molecular Biology 28(2):271-276.
- GILBERT, R.; SHINE J.; MILLER, J.; RICE, R.; RAINBOLT, C. (2004). The effect of year, cultivar, and time of harvest on sugarcane yields in Florida. Journal American Society Sugar Cane Technologists, Vol. 24.
- GOOGLE EARTH. (SF). Ubicación geográfica de los Experimentos del Estado III zafras 2014-2015 y 2015-2016. Escala indeterminada. Recuperados el 25/06/2017 desde <https://www.google.com/maps/@14.0086425,-91.1137749,1057m/data=!3m1!1e3>.
- GRIVET, L.; D'HONT, A.; DUFOUR, P.; HAMON, P.; ROQUES, D.; GLASZMANN, J.C. (1994). Comparative genome mapping of sugar cane with other species within the Andropogonae tribe. Heredity 73:500-508.
- GRIVET, L.; DANIELS, C.; GLASZMANN, J.C.; D'HONT, A. (2004). Review of recent molecular genetics evidence for sugarcane evolution and domestication. Ethnobotany Research & Applications 2:9-17.
- HADFIELD, J.D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software 33(2):1-22. URL <http://www.jstatsoft.org/v33/i02/>.
- HAIR, J.; BLACK, W.; BABIN, B.; ANDERSON, R. (2010). Multivariate Data Analysis. 7a edición. Pearson Prentice Hall. 761 p.
- HARRELL, F.E.; DUPONT, C.; Y OTROS. (2017). Hmisc: Harrell Miscellaneous. R package version 4.0-3. <https://CRAN.R-project.org/package=Hmisc>

- HENDERSON, C.R.; KEMPTHORNE, O.; SEARLE, S.R.; VON KROSIGK, C.M. (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, 15:192–218. DOI [10.2307/2527669](https://doi.org/10.2307/2527669).
- HENDERSON, C.R.; QUAAS, R.L. (1976). Multiple Trait Evaluation Using Relatives' Records. *Journal of Animal Science*. 43(6):1188-1197.
- HERNÁNDEZ, O. (1998). *Temas de análisis estadístico multivariado*. Editorial Universidad de Costa Rica. Costa Rica. 169 p.
- HILL, R.R.; ROSENBERG, J.L. (1985). Models for combining data from germoplasm evaluation trials. *Crop Sci* 25:467-470.
- HOLLAND, J.B.; NYQUIST, W.E.; CERVANTES-MARTINEZ, C.T. (2003). Estimating and Interpreting Heritability for Plant Breeding: An Update. *Plant Breeding Reviews*. 22:9-112.
- HU, B.; SHAO, J.; PALTA, M. (2006). Pseudo R² in Logistic Regression Model. *Statistica Sinica* 16:847-860.
- HU, F.; JACKSON, P.; BASFORD, K. (2008). Developing Optimal Selection Systems In Sugarcane Breeding Programs. *Proc Aust Soc Sugar Cane Technol Vol* 30:162-173.
- HUNTLEY, JW. (2011). Exploratory multivariate techniques and their utility for understanding ancient ecosystems. *In Quantifying the Evolution of Early Life*. Springer Science + Business. p. 23-48
- ICC (INSTITUTO PRIVADO DE INVESTIGACIÓN SOBRE CAMBIO CLIMÁTICO). (2011-2017). Base de datos agro-meteorológicos. Consultado el 04 de Julio 2017. Disponible en <https://redmet.icc.org.gt/>.
- INMAN-BAMBER, G.; JACKSON, P.; BONNET, G.; MORGAN, T. (2011). Have We Reached Peak CCS? *Proc Aust Soc Sugar Cane Technol Vol* 33.
- INMAN-BAMBER, G. (2014). Sugarcane Yields and Yield-Limiting Processes. *In Sugarcane: Physiology, Biochemistry, and Functional Biology*. Jhon Wiley e hijos. Iowa. p 579-600.
- ITIS. (2017). Integrated Taxonomic Information Systema Database (en línea). Washington D.C. National Museum of Natural History. Consultado 02 mayo 2017. Disponible en <https://www.itis.gov/>.
- JACKSON, P. (2005). Breeding For Improving Sugar Content in Sugarcane. *Field Crops Research*. 92:277-290.

- JACKSON, P.; HOGARTH, D.M. (1992). Genotype x environment interactions in sugarcane. I. Patterns on response across sites and crop-years in North Queensland. *Aust J Agric Res* 43:1447-1460.
- JACKSON, P.; MCRAE, T.; HOGARTH, M. (1995). Selection of sugarcane families across variable environments I. Sources of variation and an optimal selection index. *Field Crops Research* 43:109-118.
- JACKSON, P.; WEI, X.M.; ATKIN, F. (2015). Optimal selection indices in early stages trials in sugarcane breeding programs. *Proc Aust Soc Sugar Cane Technol.* 37:244-251.
- JAMES, G. An Introduction to Sugarcane. (2004). *In Sugarcane*. Blackwell Science Ltd. Oxford. p 1-19.
- JOLLIFFE, I.T.. (2002). Principal Component Analysis. Springer. Reino Unido. 487 p.
- KADIAN, S.P.; MEHLA, A.S. (2006). Correlation And Path Analysis In Sugarcane. *Indian J Agric Res* 40(1):47– 51.
- KEMPTHORNE, O.; NORDSKOG, A.W. (1959). Restricted selection indices. *Biometrics* 15:10–19.
- KHAN, I.M.; BIBI, S.; YASMIN, S.; KHATRI, A.; SEEMA, N.; ABRO, S.A. (2012). Correlation Studies Of Agronomic Traits For Higher Sugar Yield In Sugarcane. *Pak J Bot* 44(3):969-971.
- KUZNETSOVA, S.; BROCKHOFF, P.; CHRISTENSEN, R.H. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. <https://CRAN.R-project.org/package=lmerTest>
- LANDE, R.; THOMPSON, R. (1990). Efficiency of marker assisted in the improvement of quantitative traits. *Genetics* 124:743-756.
- LARRAHONDO, J.E. (1995). Calidad de la Caña de Azúcar. *In CENICAÑA. El cultivo de la caña en la zona azucarera de Colombia, Cali, CENICAÑA*, p 337-354.
- LE CUNFF, L.; GARSMEUR, O.; RABOIN, L.M.; PAUQUET, J., TELISMAR, H.; ASELVI, A.; GRIVET, L.; PHILIPPE, R.; BEGUM, D.; DEU, M.; COSTST, L.; WING, R.; GLASZMANN, J.C.; D'HONT, A. (2008). Diploid/Polyploid sintenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance Gene (*Bru 1*) in highly polyploidy sugarcane (2n~ 12x ~115). *Genetics* 180:649-660.

- LOMBARDI, G.M.; NUNES, J.A.; PARRELLA, R.A.; TEIXEIRA, D.H.; BRUZI, A.T.; DURAES, N.N.; FAGUNDES, T.G. (2015). Path analysis of agro-industrial traits in sweet sorghum. *Genetics and Molecular Research* 14 (4):16392-16402.
- LSU. (2014). Sugarcane Production Handbook. Luisiana State University. AgCenter. 85 p.
- LUO, D; GANESH, S; KOOLAARD, J. (2014). predictmeans: Calculate Predicted Means for Linear Models. R package version 0.99. <https://CRAN.R-project.org/package=predictmeans>
- MA, B.L., ZHENG, Z.M.; MORRISON, M.J. (2017). Does increasing plant population density alter sugar yield in high stalk-sugar maize hybrids? *Crop & Pasture Science* 68:, 1–10.
- MANNERS, JM. (2011). Functional genomics of sugarcane. *Advances in Botanical Reseach*. 60:89-168.
- MANNING, C. (2007). Logistic Regression (with R) (en línea), Stanford. 15 p. consultado 04 mayo 2017. Disponible en <https://nlp.stanford.edu/manning/courses/ling289/logistic.pdf>.
- MASSY, W. (1965). Principal Components Regression in Exploratory Statistical Reseach. *Journal or the American Statistical Association*. 60:234-256.
- MEDINA, E. (2003). Modelos de Elección Discreta. Consultado el 25 Octubre 2016. Disponible en www.eva.medinaam.es
- MELGAR, M. (2012). Desarrollo Tecnológico de la Agroindustria Azucarera y Perspectivas. En: *El cultivo de la caña de azúcar en Guatemala*. Melgar, M; Meneses, A; Orozco, H; Pérez, O; Espinosa, R. Eds. Guatemala, CENGICAÑA. Artemis Edinter. p 1-31.
- MENDEL, G. (1866). Experiments in plant hybridization. *Abhandlungen*, 3–47.
- MENDENHAL, W.; BEAVER, R. (1991). *Introduction to Probability and Statistics*. International Thomson Publishing. 8a edición. 848 p.
- MILLER, J.D.; JAMES, N.; LYRENE, P. (1978). Selection Indices in Sugarcane. *Crop Science* 18(3):369-372.
- MING, R.; LIU, P.; MOORE, P.; IRVINE, J.; PATERSON, H. (2001). QTL Analysis in a Complex Autopolyploid: Genetic Control of Sugar Content in Sugarcane. *Genome Research* 11:2075–2084.
- MING, R.; MOORE, P.H.; WU, K.K.; D'HONT, A., GLASZMANN, J.C.; TEW, T.L.; MIRKOV, T.E.; DA SILVA, J.; JIFON, J.; RAI, M.; SCHNELL, R.J.;

- BRUMBLEY, S.M.; LAKSHMANAN, P.; COMSTOCK, J.C.; PATERSON, A.H. (2006). Sugarcane Improvement through Breeding and Biotechnology. *In* (J Janick ed.). *Plant Breeding Reviews* 27: 15-118.
- MONTEIRO, L.A.; SENTELHAS, P.C. (2014). Potential and Actual Sugarcane Yields in Southern Brazil as a Function of Climate Conditions and Crop Management. *Sugar Tech* 16(3):264-276.
- MONTGOMERY, D. (1991). *Diseño y Análisis de Experimentos*. Trad. J Delgado. México. Grupo Editorial Iberoamérica. 589 p.
- MOORE, P.H.; BERDING, N. (2014). Flowering. *In* *Sugarcane: Physiology, Biochemistry, and Functional Biology*. Jhon Wiley e hijos. Iowa. p 379-410.
- MOORE, P.H.; PATERSON, A.H.; TEW, T. (2014). Sugarcane: the crop, the plant and domestication. *In* *Sugarcane: Physiology, Biochemistry, and Functional Biology*. Jhon Wiley e hijos. Iowa. p 1-17.
- NAGARAJAN, R.; ALARMELU, S.; SHANTHI, R.M.. (2006). Variability For Yield And Quality Attributes *In* Interspecific Progenies Of *Saccharum* sp. Second National Plant Breeding Congress 2006. Plant Breeding in Post Genomics Era. India.
- NELDER, JA. (1977). A Reformulation of Linear Models. *J. R. Statist. Soc. A*. 140: 48-77.
- NIA, M.S.; EMAMI, H.; HONARNEJAD, R.; ESFAHANI, M. (2011). Comparing Economic Coefficients to Select the Optimum Selectin Index in Peanuts. *American-Eurasian J Agric & Environ Sci* 10(6):972-977.
- OECD/FAO. (2016), *OECD-FAO Agricultural Outlook 2016-2025*, OECD Publishing, Paris. http://dx.doi.org/10.1787/agr_outlook-2016-en
- ONG'ALA, J.; MWANGA, D.; OUANI, F. (2016). On the Use of Principal Component Analysis in Sugarcane Clone Selection. *Journal of the Indian Society of Agricultural Statistics* 70(1):33-39.
- OROZCO, H.; BUC, R. (2016). Censo de variedades de caña de azúcar en Guatemala, zafra 2015-2016. *In*, Memoria. Presentación de resultados de investigación. Zafra 2015-2016. CENGICAÑA, Guatemala. En discos compactos. 463 p.
- OROZCO, H.; QUEMÉ, J.; OVALLE, W.; ROSALES-LONGO, F. (2012). Mejoramiento genético de la caña de azúcar. *In* *El cultivo de la caña de azúcar en Guatemala*. Melgar, M; Meneses, A; Orozco, H; Pérez, O; Espinosa, R. Eds. Guatemala, CENGICAÑA. Artemis Edinter. p 45-79.

- OROZCO, H.; BUC, R.; OVALLE, W.; QUEMÉ, J.; AZAÑÓN, V.; GÓMEZ, V.; PAZ, V. (2016). Selección de variedades para uso comercial de la séptima prueba semicomercial en la agroindustria azucarera guatemalteca. *In* Memoria. Presentación de resultados de investigación Zafra 2015-2016. CENGICAÑA, Guatemala. En discos compactos. 462 p.
- ORTIZ, S.; VALDÉS, M.P.; VALLEJO, F.A.; BAENA, D. (2015). Genetic Correlations and Path Analysis in Butternut Squash *Cucurbita moschata* Duch. *Rev.Fac.Nal.Agr.Medellín* 68(1): 7399-7409
- PASTINA, M. M.; MALOSETTI, M.; GAZAFFI, R.; MOLLINARI, M.; MARGARIDO, G.R.A.; OLIVEIRA, K.M.; GARCIA, A.A.F. (2012). A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. *TAG. Theoretical and Applied Genetics*. 124(5), 835–849. <http://doi.org/10.1007/s00122-011-1748-8>
- PÉREZ, O.; HERNÁNDEZ, F.; MARTÍNEZ, C.; AZAÑÓN, V.; TAYÚN, J.; TUCHÁN, L.; PAZ, V.; JAVIER, A. (2016). Respuesta de nuevas variedades de caña de azúcar al nitrógeno en la zona cañera de Guatemala. *In*, Memoria. Presentación de resultados de investigación. Zafra 2015-2016 (disco compacto). Guatemala, CENGICAÑA 1 disco compacto.
- PÉREZ-ELIZALDE, S.; CERÓN-ROJAS, J.J.; CROSSA, J.; FLEURY, D.; ALVARADO, G. (2014). *Rindsel*: An R Package for Phenotypic and Molecular Selection Indices Used in Plant Breeding. *In* Crop Breeding: Methods and Protocols, *Methods in Molecular Biology*, vol. 1145. Springer Science and Business Media New York.
- PESEK, J.; BAKER, R.J. (1970). An Application Of Index Selection To The Improvement Of Self Pollinated Species. *Can J Plant Sci* 50:267-276.
- PIEPHO, H.; MÖHRING, J.; MELCHINGER, A.; BÜCHSE, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* (2008) 161:209–228.
- PILLAI, S.; ETHIRAJAN, A.S. (1993). SELECTION INDICES FOR SUGARCANE IMPROVEMENT AT THREE STAGES OF SELECTION. *EUPHYTICA* 1(2):155-159.
- PIPERIDIS, N.; CHEN, J.; DENG, H.; WANG, L.; JACKSON, P.; PIPERIDIS, G. (2010). GISH characterization of *Erianthus arundinaceus* chromosomes in three generations of sugarcane in intergenérico hybrids. *Genome* 53:331-336.
- PURSEGLOVE J. (1972). Sugarcane. *In* Tropical Crops Monocotyledons. Longman Group. P 214-256.

- QUEMÉ, J.L.; MELGAR, M.F. (2016). La Remolacha Azucarera En Francia Y La Caña De Azúcar En Guatemala. *In Memoria. Presentación de resultados de investigación Zafra 2015-2016. CENGICAÑA, Guatemala. En discos compactos. 462 p.*
- QUEMÉ, J.L.; MOLINA, L.; MELGAR, M. (2005). Analysis of genetic similarity among 48 sugarcane varieties using microsatellite DNA sequences. *Proc. ISSCT (25):592-596.*
- QUEMÉ, J.L.; CROSSA, J.; OROZCO, H.; MELGAR, M. (2007). Analysis of genotype-by-environment interaction for sugarcane using the sites regression model (SREG). *Proc. Int. Soc. Sugar Cane Technol. Vol. 26.*
- QUEME, J.; OROZCO, H.; MELGAR, M. (2010). GGE biplot analysis used to evaluate cane yield of sugarcane (*Saccharum* spp.) cultivars across sites and crop cycles. *Proc. Int. Soc. Sugar Cane Technol. Vol 27.*
- QUEMÉ, J.L.; OROZCO, H.; SALAZAR, A. (2016). Resultados Relacionados Con El Recurso Genético Y Cruzamientos Del Programa de Variedades De Cengicaña En La Zafra 2015-2016. *In Memoria. Presentación de resultados de investigación Zafra 2015-2016. CENGICAÑA, Guatemala. En discos compactos. 462 p.*
- R CORE TEAM (2017). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- RAMBURAN, S.; ZHOU, M.; LABUSCHAGNE, M. (2011). Interpretation of genotype \times environment interactions of sugarcane: Identifying significant environmental factors. *Field Crops Res, 124:3-20.*
- RENCER, AC. (2002). *Methods of multivariate analysis.* Wiley-Interscience. Jhon Wiley and Sons. Canada. 727 p.
- ROBINSON, GK. (1991). That BLUP is a Good Thing: The Estimation of Random Effects. *Statistical Science, 6(1):15-32.*
- RODRÍGUEZ, R.; PUCHADES, Y.; BERNAL, N.; JORGE, H.; TAMAYO, M.; CARABALLOSO, V. (2013). Location contribution of multi-environment sugarcane trials determined via GGE biplot analysis. *Proc. Int. Soc. Sugar Cane Technol., Vol. 28.*
- ROSALES-LONGO, F. (2011). Aspectos generales de la citogenética de la caña de azúcar. *In, Memoria. Presentación de resultados de investigación. Zafra 2011-2012. CENGICAÑA, Guatemala. p 103-106.*
- ROSALES-LONGO, F.; OVALLE, W.; GARCÍA, B.; GONZÁLEZ, H.; CATALÁN, M.; LIMA, L.; PAZ, V. (2016a). Evaluación y selección de variedades de caña de azúcar

- en el estado II durante la zafra 2015-2016, en Guatemala. *In Memoria*. Presentación de resultados de investigación Zafra 2015-2016. Guatemala. En discos compactos. 462 p.
- ROSALES-LONGO, F.; OVALLE, W.; GARCÍA, B.; GONZÁLEZ, H.; CATALÁN, M.; GARCÍA, S.; LIMA, L.; PAZ, V.; VILLATORO, B. (2016b). Evaluación y selección de variedades de caña de azúcar en el estado III durante la zafra 2015-2016, en Guatemala. *In Memoria*. Presentación de resultados de investigación Zafra 2015-2016. Guatemala. En discos compactos. 462 p.
- SAGE, R.; PEIXOTO, M.; SAGE, TL. (2014). Photosynthesis in Sugarcane. *In Sugarcane: Physiology, Biochemistry, and Functional Biology*. Wiley, Oxford, Reino Unido. p. 121:154.
- SALAZAR, F.; VICTORIA, J. (2013). Effect of climate variables on the selection of sugarcane varieties for the Cauca river valley of Colombia. *Proc. Int. Soc. Sugar Cane Technol*, Vol. 28.
- SÁNCHEZ, G. (2013). PLS Path Moedeling With R. Trowchez Editions. Berkeley. 235 p.
- SCHOLERKE, B.; CROWLEY, J.; DI COOK; BRIATTE, F.; MARBACH, M.; THOEN, E.; ELBERG A.; LARMARANGE, J. (2016). GGally: Extension to 'ggplot2'. R package version 1.3.0. <https://CRAN.R-project.org/package=GGally>
- SEMAGN, K.; BJØRNSTAD Å.; XU, Y. (2010). The genetic dissection of quantitative traits in crops. *Electronic Journal of Biotechnology* 13(5). <http://dx.doi.org/10.2225/vol13-issue5-fulltext-14>
- SHANTI, P.M.; REDY, K.H.P.; HEMANTH, K.M.; RAJARAJESWARI, V.; MOHAN, N.G.; NARASIMHULU, R.; RUPESH, K.R.; MOHAN, R.D. (2013). Selection Indices For Cane Yield In Sugarcane (*Saccharum* spp.). *Indian Journal of Plant Sciences*. 2(3):77-81.
- SILVA, E.; CASTILLO, F.; MOLINA, J.; BENITEZ, I.; SANTA-CRUZ, A; CASTILLO, R. (2011). Selección de progenitores, varianzas genéticas y heredabilidad para acumulación temprana de sacarosa en caña de azúcar. *Rev Fitotec Mex*. 34(2):107-114.
- SIMEÃO, R.M.; STURION, J.A.; RESENDE, M.D.; FERNANDES, J.S.; NEIVERTH, D.D.; ULBRICH, A.L. (2002). Avaliação genética em erva-mate pelo procedimento BLUP individual multivariado sob interação genótipo x ambiente. *Pesq. agropec. bras.*, Brasília. 37(11):1589-1596.

- SIMON, S.; HEMAPRABHA, G. (2010). Identification of two new drought specific candidate genes in sugarcane (*Saccharum* spp.). *Electronic journal of plant breeding* 1(4):1164-1170.
- SINGH R.K., MISRHA S.K., SINGH S.P., MISHRA N., SHARMA M.L. (2010). Evaluation of microsatellite markers for genetic diversity analysis among sugarcane species and commercial hybrids. *Australian Journal of Crop Science* 4(2):116-125.
- SINGH, S.P.; KHAN, A.Q. (1998). Selection Indices For Cane Yield In Sugarcane. *Indian J Genet* 58(3): 353-357.
- SKINNER, J.C.; HOGARTH, M.; WU, K. (1987). Selection methods, criteria, and indices. *In Sugarcane improvement through breeding*. Elsevier. p 409-453.
- SMITH, H.F. (1936). A discriminant function for plant selection. *Annual Eugenics* 7: 40-250.
- SPIEGEL, M.R. (1987). *Probabilidad y Estadística. Serie Schaum*. McGraw-Hill. México. 372 p.
- SREENIVASAN, T.; AHLOOWALIA, B.; HEINZ, D. (1987). Cytogenetics. *In Sugarcane improvement through breeding*. Ed. Heinz. Elsevier Science. p 211-253.
- STEEL, R.G.; TORRIE, J.H. (1986). *Bioestadística: Principios y Procedimientos*. McGraw-Hill. México. 622 p.
- TAHIR, M.; KHALIL, I.H.; MCCORD, P.H.; GLAZ, B. (2014). Character Association and Selection Indices in Sugarcane. *American Journal of Experimental Agriculture* 4(3):336-348.
- TAYLOR, J.; VERBYLA, A. (2011). R Package wgaim: QTL Analysis in Bi-Parental Populations Using Linear Mixed Models. *Journal of Statistical Software* 40(7):1-17.
- TENA, E.; MEKBIB, F.; AYANA, A. (2016a). Genetic Diversity of Quantitative Traits of Sugarcane Genotypes in Ethiopia. *American Journal of Plant Science*. 7:1498-1520.
- TENA, E.; MEKBIB, F.; AYANA, A. (2016b). Heritability and Correlation among Sugarcane (*Saccharum* spp.) Yield and Some Agronomic and Sugar Quality Traits in Ethiopia. *American Journal of Plant Sciences* 7:1453-1477.
- THOMPSON, R. (2008). Estimation of quantitative genetic parameters. *Proc. R. Soc.* 275:679-686.

- VAN DER MERWE, M.J; BOTHA, F.C. (2014). Respiration as a Competitive Sink for Sucrose Accumulation in Sugarcane Culm. *In* Sugarcane: Physiology, Biochemistry, and Functional Biology. Wiley, Oxford, Reino Unido. p. 121-154.
- VAN DER WERF, J. (2002). Mixed Models for Genetic Analysis. Universidad de New England, Armidale, Australia. 27 p. Consultado 02 junio 2017. Disponible en https://jvanderw.une.edu.au/Mixed_Models_for_Genetic_analysis.pdf.
- VENABLES, W.N; RIPLEY, B.D. (2002). Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
- VILLATORO, B.; MENESES, A.A. (2013). Manejo agronómico por ambiente (MAPA) enfoque para aumentar la productividad de la agroindustria azucarera de Guatemala. *In*, Memoria. Presentación de resultados de investigación. Zafra 2012-2013 (disco compacto). Guatemala, CENGICAN 1 disco compacto.
- VIVEROS, C.; BAENA, D.; VILLARREAL, F.; LÓPEZ, L.O.; VICTORIA, J. (2015). Traits of the sugar cane associated with tons of cane per hectare and sucrose (% cane). *Acta Agronómica*. 64(3): 252-256.
- WAGNER, R. (2007). Historia de la Caña de Azúcar en Guatemala. Galería. Fundación G & T/Asociación de Azucareros de Guatemala. Guatemala.
- WATT, D.; MCCORMICK, A.J.; CRAMER, M.D. (2014). Source and Sink Physiology. *In* Sugarcane: Physiology, Biochemistry, and Functional Biology. Wiley-Blackwell. Iowa. p 483-520.
- WEI, T.; SIMKO, V. (2016). corrplot: Visualization of a Correlation Matrix. R package version 0.77. <https://CRAN.R-project.org/package=corrplot>
- WICKHAM, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21(12), 1-20. URL <http://www.jstatsoft.org/v21/i12/>
- WICKHAM, H. (2009). Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
- WICKHAM H.; CHANG, W. (2017). devtools: Tools to Make Developing R Packages Easier. R package version 1.13.1. <https://CRAN.R-project.org/package=devtools>
- WILLIAMS, J.S. (1962). The Evaluation Of A Selection Index. *Biometrics* 18:375-393.





ANEXOS



ANEXOS

ANEXO 1. LISTADO DE GENOTIPOS EN CADA EXPERIMENTO. LAS CELDAS SOMBREADAS EN CELESTE, CORRESPONDEN A LAS VARIEDADES TESTIGOS EN CADA EXPERIMENTO.

No.	Ensayo	Variedad	Progenitor Femenino	Progenitor(es) Masculino(s)
1	litoral_flor	CG12-534048	CP89-1431	CP63-588
2	litoral_flor	CG12-333035	CP72-1210	CC85-92
3	litoral_flor	CG12-627018	CP73-1547	SP79-2233
4	litoral_flor	CG12-339029	CP63-588	Q96
5	litoral_flor	CG12-626004	SP79-2233	CP73-1547/SP70-1284
6	litoral_flor	CG12-610094	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
7	litoral_flor	CG12-610036	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
8	litoral_flor	CG12-059078	CP99-1944	CP03-2420
9	litoral_flor	CG12-008055	CP62-374	CP81-1425
10	litoral_flor	CG12-533018	CP89-1431	CP63-588
11	litoral_flor	CG12-404004	SP79-2233	CP72-2086
12	litoral_flor	CG12-MSING#52	?	?
13	litoral_flor	CG12-638003	10P20602	BR9423
14	litoral_flor	CG12-034060	CP02-1143	CP93-1361/CP78-1610
15	litoral_flor	CG12-508001	CP70-1133	CG96-01
16	litoral_flor	CG12-045002	CP73-1547	CG00-044
17	litoral_flor	CG12-115002	C87-51	CP03-1812
18	litoral_flor	CG12-328001	CP65-357	CP88-1508/CG98-46
19	litoral_flor	CG12-055003	CP90-1017	CP91-2234/CP92-1314
20	litoral_flor	CG12-253003	SP71-6180	CP70-1133
21	litoral_flor	CG12-318003	CB38-22	CG98-41/CC85-63
22	litoral_flor	CG12-015002	L79-43	CP86-1830
23	litoral_flor	CG12-032004	CP02-1143	CP27-139
24	litoral_flor	CG12-134009	CG00-102	CP88-1165/CP73-1547
25	litoral_flor	CG12-621061	CP88-1508	L80-2/CP57-603
26	litoral_flor	CG12-051044	CP75-1091	CP77-1055
27	litoral_flor	CG12-627056	CP73-1547	SP79-2233
28	litoral_flor	CG12-338021	CP63-588	CP89-1431
29	litoral_flor	CG12-338135	CP63-588	CP89-1431
30	litoral_flor	CG12-621036	CP88-1508	L80-2/CP57-603
31	litoral_flor	CG12-036004	CP02-1143	CP81-1384
32	litoral_flor	CG12-326003	CG96-01	B74132

33	litoral_flor	CG12-036096	CP02-1143	CP81-1384
34	litoral_flor	CG12-094002	CP72-1210	CC85-92/CP82-2043
35	litoral_flor	CG12-360009	CP72-1210	B74132/CP89-2143
36	litoral_flor	CG12-055001	CP90-1017	CP91-2234/CP92-1314
37	litoral_flor	CG12-611002	CP70-1133	CP68-1067/CP70-321/CP72-1312
38	litoral_flor	CG12-325006	CG96-01	B74132
39	litoral_flor	CG12-281001	L80-42	V71-51
40	litoral_flor	CG12-564002	CC85-92	CP65-357
41	litoral_flor	CG12-105001	CP81-1425	CP43-47/Ja55-488
42	litoral_flor	CG12-311014	CG96-01	CP70-1133
43	litoral_flor	CG12-613003	CP82-2043	CP81-1384/SP79-2233/CP33-224
44	litoral_flor	CG12-606012	CC85-63	CC84-75/CB38-22/B4362/Akoki
45	litoral_flor	CG12-059028	CP99-1944	CP03-2420
46	litoral_flor	CG12-325003	CG96-01	B74132
47	litoral_flor	CG12-059076	CP99-1944	CP03-2420
48	litoral_flor	CG12-116008	CP80-1827	CP03-1812
49	litoral_flor	CG12-325052	CG96-01	B74132
50	litoral_flor	CG12-325046	CG96-01	B74132
51	litoral_flor	CG12-333051	CP72-1210	CC85-92
52	litoral_flor	CG12-621016	CP88-1508	L80-2/CP57-603
53	litoral_flor	CG12-065014	RB739331	CP02-2188
54	litoral_flor	CG12-606013	CC85-63	CC84-75/CB38-22/B4362/Akoki
55	litoral_flor	CG12-332009	CP72-1210	CP82-2043
56	litoral_flor	CG12-531023	CP89-2143	CP73-1547
57	litoral_flor	CG12-387001	CP65-357	RB732908
58	litoral_flor	CG12-353003	CP63-588	CP88-1508
59	litoral_flor	CG12-014043	CP29-103	Akoki
60	litoral_flor	CG12-135002	CG00-102	CP02-1143/CG98-46
61	litoral_flor	CG12-324191	CG96-01	B7306
62	litoral_flor	CG12-318023	CB38-22	CG98-41/CC85-63
63	litoral_flor	CG12-324175	CG96-01	B7306
64	litoral_flor	CG12-324199	CG96-01	B7306
65	litoral_flor	CG12-333011	CP72-1210	CC85-92
66	litoral_flor	CG12-338029	CP63-588	CP89-1431
67	litoral_flor	CG12-326007	CG96-01	B74132
68	litoral_flor	CG12-627031	CP73-1547	SP79-2233
69	litoral_flor	CG12-333064	CP72-1210	CC85-92
70	litoral_flor	CG12-629008	CG96-01	CP65-357
71	litoral_flor	CP88-1165	CL61-620	CP81-1302
72	litoral_flor	CG12-326012	CG96-01	B74132
73	litoral_flor	CG12-610029	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
74	litoral_flor	CG12-326016	CG96-01	B74132

75	litoral_flor	CG12-333002	CP72-1210	CC85-92
76	litoral_flor	CG12-629041	CG96-01	CP65-357
77	litoral_flor	CG12-326036	CG96-01	B74132
78	litoral_flor	CG12-621058	CP88-1508	L80-2/CP57-603
79	litoral_flor	CG12-627044	CP73-1547	SP79-2233
80	litoral_flor	CG12-324082	CG96-01	B7306
81	litoral_flor	CG12-621060	CP88-1508	L80-2/CP57-603
82	litoral_flor	CG12-338070	CP63-588	CP89-1431
83	litoral_flor	CG12-333034	CP72-1210	CC85-92
84	litoral_flor	CG12-339050	CP63-588	Q96
85	litoral_flor	CG12-338050	CP63-588	CP89-1431
86	litoral_flor	CG12-621044	CP88-1508	L80-2/CP57-603
87	litoral_flor	CG12-621055	CP88-1508	L80-2/CP57-603
88	litoral_flor	CG12-339013	CP63-588	Q96
89	litoral_flor	CG12-584034	Mex57-29	CP65-357/CP88-1508
90	litoral_flor	CG12-324058	CG96-01	B7306
91	litoral_flor	CG12-629062	CG96-01	CP65-357
92	litoral_flor	CG12-610031	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
93	litoral_flor	CG12-627023	CP73-1547	SP79-2233
94	litoral_flor	CG12-621027	CP88-1508	L80-2/CP57-603
95	litoral_flor	CG12-324031	CG96-01	B7306
96	litoral_flor	CG12-326001	CG96-01	B74132
97	litoral_flor	CG12-621048	CP88-1508	L80-2/CP57-603
98	litoral_flor	CG12-MSING#36	?	?
99	litoral_flor	CG12-338044	CP63-588	CP89-1431
100	litoral_flor	CG12-286003	CB38-22	CP82-2043/CP70-1133
101	litoral_flor	CG12-360004	CP72-1210	B74132/CP89-2143
102	litoral_flor	CG12-339002	CP63-588	Q96
103	litoral_flor	CG12-339080	CP63-588	Q96
104	litoral_flor	CG12-606034	CC85-63	CC84-75/CB38-22/B4362/Akoki
105	litoral_flor	CG12-620049	IJ76-319	Hinahina/IJ76-441/CP91-2249
106	litoral_flor	CG12-534070	CP89-1431	CP63-588
107	litoral_flor	CG12-339136	CP63-588	Q96
108	litoral_flor	CG12-034045	CP02-1143	CP93-1361/CP78-1610
109	litoral_flor	CG12-339042	CP63-588	Q96
110	litoral_flor	CG12-620040	IJ76-319	Hinahina/IJ76-441/CP91-2249
111	litoral_flor	CG12-317005	CB38-22	CG98-41/CC85-63
112	litoral_flor	CG12-534069	CP89-1431	CP63-588
113	litoral_flor	CG12-339071	CP63-588	Q96
114	litoral_flor	CG12-534083	CP89-1431	CP63-588
115	litoral_flor	CG12-534044	CP89-1431	CP63-588
116	litoral_flor	CG12-014008	CP29-103	Akoki

117	litoral_flor	CG12-324206	CG96-01	B7306
118	litoral_flor	CG12-326002	CG96-01	B74132
119	litoral_flor	CG12-339103	CP63-588	Q96
120	litoral_flor	CG12-136009	CP91-1696	CP70-1133
121	litoral_flor	CG12-136016	CP91-1696	CP70-1133
122	litoral_flor	CG12-312003	CG96-01	CP70-1133
123	litoral_flor	CG12-040023	CP29-103	CP77-1055/HoCP01-523
124	litoral_flor	CG12-534046	CP89-1431	CP63-588
125	litoral_flor	CG12-596006	CP89-2143	CG96-01
126	litoral_flor	CG12-621002	CP88-1508	L80-2/CP57-603
127	litoral_flor	CG12-621013	CP88-1508	L80-2/CP57-603
128	litoral_flor	CG12-621037	CP88-1508	L80-2/CP57-603
129	litoral_flor	CG12-621064	CP88-1508	L80-2/CP57-603
130	litoral_flor	CG12-621065	CP88-1508	L80-2/CP57-603
131	litoral_flor	CG12-621067	CP88-1508	L80-2/CP57-603
132	litoral_flor	CG12-627080	CP73-1547	SP79-2233
133	litoral_flor	CG12-629059	CG96-01	CP65-357
134	litoral_flor	CG98-46	CP56-59	CP57-603
135	litoral_flor	CGB12-635007	10P20462	BR9423
136	litoral_flor	CGB12-635058	10P20462	BR9423
137	litoral_flor	CP72-2086	CP62-374	CP63-588
138	litoral_flor	CP73-1547	CP66-1043	CP56-63
1	litoral_no_flor	CG12-330012	Eros	CP75-1082
2	litoral_no_flor	CG12-564002	CC85-92	CP65-357
3	litoral_no_flor	CG12-606129	CC85-63	CC84-75/CB38-22/B4362/Akoki
4	litoral_no_flor	CG12-324085	CG96-01	B7306
5	litoral_no_flor	CG12-333040	CP72-1210	CC85-92
6	litoral_no_flor	CG12-105	CP63-588	CP89-1431
7	litoral_no_flor	CG12-103	CP91-1696	CP70-1133
8	litoral_no_flor	CG12-629011	CG96-01	CP65-357
9	litoral_no_flor	CG12-338092	CP63-588	CP89-1431
10	litoral_no_flor	CG12-333044	CP72-1210	CC85-92
11	litoral_no_flor	CG12-317039	CB38-22	CG98-41/CC85-63
12	litoral_no_flor	CG12-606033	CC85-63	CC84-75/CB38-22/B4362/Akoki
13	litoral_no_flor	CG12-568025	CC85-63	CP63-588/CP89-2143
14	litoral_no_flor	CG12-339052	CP63-588	Q96
15	litoral_no_flor	CG12-326030	CG96-01	B74132
16	litoral_no_flor	CG12-MSING#12	?	?
17	litoral_no_flor	CG12-330008	Eros	CP75-1082
18	litoral_no_flor	CG12-034026	CP02-1143	CP93-1361/CP78-1610
19	litoral_no_flor	CG12-317022	CB38-22	CG98-41/CC85-63
20	litoral_no_flor	CG12-014012	CP29-103	Akoki
21	litoral_no_flor	CG12-113	CP89-1431	CP63-588

22	litoral_no_flor	CG12-102	CP02-1143	CP27-139
23	litoral_no_flor	CG12-060023	CP99-1944	HoCP02-610
24	litoral_no_flor	CG12-112	CP88-1508	Mex79-431
25	litoral_no_flor	CG12-059010	CP99-1944	CP03-2420
26	litoral_no_flor	CG12-526031	CP75-1082	Eros
27	litoral_no_flor	CG12-030014	CP70-1133	CP77-1088
28	litoral_no_flor	CG12-534017	CP89-1431	CP63-588
29	litoral_no_flor	CG12-339033	CP63-588	Q96
30	litoral_no_flor	CG12-338004	CP63-588	CP89-1431
31	litoral_no_flor	CG12-339039	CP63-588	Q96
32	litoral_no_flor	CG12-338119	CP63-588	CP89-1431
33	litoral_no_flor	CG12-548018	CC85-92	CP63-588
34	litoral_no_flor	CG12-338087	CP63-588	CP89-1431
35	litoral_no_flor	CG12-324094	CG96-01	B7306
36	litoral_no_flor	CG12-338115	CP63-588	CP89-1431
37	litoral_no_flor	CG12-338013	CP63-588	CP89-1431
38	litoral_no_flor	CG12-339010	CP63-588	Q96
39	litoral_no_flor	CG12-135043	CG00-102	CP02-1143/CG98-46
40	litoral_no_flor	CG12-114	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
41	litoral_no_flor	CG12-136018	CP91-1696	CP70-1133
42	litoral_no_flor	CG12-311028	CB38-22	CG98-41/CC85-63
43	litoral_no_flor	CG12-324083	CG96-01	B7306
44	litoral_no_flor	CG12-627006	CP73-1547	SP79-2233
45	litoral_no_flor	CG12-311019	CG96-01	CP70-1133
46	litoral_no_flor	CG12-MSING#66	?	?
47	litoral_no_flor	CG12-386026	CP65-357	RB732908
48	litoral_no_flor	CG12-523061	CP88-1508	CP65-357/CG98-46
49	litoral_no_flor	CG12-638006	10P20602	BR9423
50	litoral_no_flor	CG12-324227	CG96-01	B7306
51	litoral_no_flor	CG12-616017	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
52	litoral_no_flor	CG12-324204	CG96-01	B7306
53	litoral_no_flor	CG12-339105	CP63-588	Q96
54	litoral_no_flor	CG12-065010	RB739331	CP02-2188
55	litoral_no_flor	CG12-311020	CG96-01	CP70-1133
56	litoral_no_flor	CG12-317012	CB38-22	CG98-41/CC85-63
57	litoral_no_flor	CG12-101	CP70-1133	CP01-2409/CP93-1361
58	litoral_no_flor	CG12-036023	CP02-1143	CP81-1384
59	litoral_no_flor	CG12-337009	CP63-588	CP89-1431
60	litoral_no_flor	CG12-532008	CP88-1508	Mex79-431
61	litoral_no_flor	CG12-338009	CP63-588	CP89-1431
62	litoral_no_flor	CG12-318005	CB38-22	CG98-41/CC85-63
63	litoral_no_flor	CG12-326047	CG96-01	B74132
64	litoral_no_flor	CG12-339117	CP63-588	Q96
65	litoral_no_flor	CG12-339025	CP63-588	Q96

66	litoral_no_flor	CG12-620015	IJ76-319	Hinahina/IJ76-441/CP91-2249
67	litoral_no_flor	CG12-115	10P20410	WI93904
68	litoral_no_flor	CG12-373005	L76-331	CP89-2143
69	litoral_no_flor	CG12-606112	CC85-63	CC84-75/CB38-22/B4362/Akoki
70	litoral_no_flor	CG12-039004	CG98-36	HoCP01-523
71	litoral_no_flor	CG12-020001	CP92-1607	B7306
72	litoral_no_flor	CG12-036037	CP02-1143	CP81-1384
73	litoral_no_flor	CG12-330063	Eros	CP75-1082
74	litoral_no_flor	CG12-626035	SP79-2233	CP73-1547/SP70-1284
75	litoral_no_flor	CG12-532007	CP88-1508	Mex79-431
76	litoral_no_flor	CG12-339018	CP63-588	Q96
77	litoral_no_flor	CG12-568021	CC85-63	CP63-588/CP89-2143
78	litoral_no_flor	CG12-606095	CC85-63	CC84-75/CB38-22/B4362/Akoki
79	litoral_no_flor	CG12-018003	CP03-1812	CP80-1827
80	litoral_no_flor	CG12-630003	10P60724	BT84118
81	litoral_no_flor	CG12-111	DB5/55	CG96-01/Q96
82	litoral_no_flor	CG12-106001	Ja55-488	CP43-47/CP81-1425
83	litoral_no_flor	CG12-132001	CG00-044	CL61-620
84	litoral_no_flor	CG12-531002	CP89-2143	CP73-1547
85	litoral_no_flor	CG12-239002	CP73-1547	CP81-1384
86	litoral_no_flor	CG12-108	CP81-1384	CP88-1165
87	litoral_no_flor	CG12-109	CP81-1384	CP88-1165
88	litoral_no_flor	SP79-1287	?	?
89	litoral_no_flor	CG12-039008	CG98-36	HoCP01-523
90	litoral_no_flor	CG98-10	PR87-2078	PR87-2073
91	litoral_no_flor	CG12-581002	CP89-2143	SP91-2074
92	litoral_no_flor	CG12-239005	Ja64-20	CP91-2161
93	litoral_no_flor	CP88-1165	CL61-620	CP81-1302
94	litoral_no_flor	CG12-353005	CP63-588	CP88-1508
95	litoral_no_flor	CG12-568011	CC85-63	CP63-588/CP89-2143
96	litoral_no_flor	CG12-319018	CB38-22	CG98-41/CC85-63
97	litoral_no_flor	CG12-624002	PGM89-968	PGM89-118/POJ
98	litoral_no_flor	CG12-204003	CP62-374	B70121
99	litoral_no_flor	CP72-2086	CP62-374	CP63-588
100	litoral_no_flor	CG12-005031	L56-20	CP81-1425
101	litoral_no_flor	CG12-606014	CC85-63	CC84-75/CB38-22/B4362/Akoki
102	litoral_no_flor	CG12-110	Q96	CG96-01/DB5/55
103	litoral_no_flor	CG12-104	CG96-01	CP70-1133
104	litoral_no_flor	CG12-276003	CP65-357	CG98-46
105	litoral_no_flor	CG12-370003	CP63-588	CC85-63/CP89-2143
106	litoral_no_flor	CG12-057015	CP99-1707	CP47-193
107	litoral_no_flor	CG12-107	CP65-357	RB732908
108	litoral_no_flor	CG12-395006	CP33-224	CP72-2086
109	litoral_no_flor	CG12-512004	CG98-41	CB38-22/CC85-63

110	litoral_no_flor	CG12-368002	CG96-01	CP89-2143
111	litoral_no_flor	CG12-317021	CB38-22	CG98-41/CC85-63
112	litoral_no_flor	CG12-116	CC85-63	CC84-75/CB38-22/B4362/Akoki
113	litoral_no_flor	CG12-229001	CP92-1401	CP72-1312
114	litoral_no_flor	CG12-106	CP65-357	CC85-92
115	litoral_no_flor	CG12-060016	CP99-1944	HoCP02-610
116	litoral_no_flor	SP71-6161	?	?
117	litoral_no_flor	CG12-317031	CB38-22	CG98-41/CC85-63
1	medio_flor	CG12-627036	CP73-1547	SP79-2233
2	medio_flor	CG12-621007	CP88-1508	L80-2/CP57-603
3	medio_flor	CG12-035012	CP02-1143	CP78-1610
4	medio_flor	CG12-210004	CP86-1209	CP81-1384
5	medio_flor	CG12-014007	CP29-103	Akoki
6	medio_flor	CG12-620011	IJ76-319	Hinahina/IJ76-441/CP91-2249
7	medio_flor	CG12-324105	CG96-01	B7306
8	medio_flor	CG12-534090	CP89-1431	CP63-588
9	medio_flor	CG12-629056	CG96-01	CP65-357
10	medio_flor	CG12-325020	CG96-01	B74132
11	medio_flor	CG12-333017	CP72-1210	CC85-92
12	medio_flor	CG12-146003	HoCP02-610	CP99-1944
13	medio_flor	CG12-325061	CG96-01	B74132
14	medio_flor	CG12-041003	L80-3	CP78-1610
15	medio_flor	CG12-610019	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
16	medio_flor	CG12-523009	CP88-1508	CP65-357/CG98-46
17	medio_flor	CG12-260004	SP79-2233	CP73-1547/SP70-1284
18	medio_flor	CG12-339356	CP63-588	Q96
19	medio_flor	CG12-326014	CG96-01	B74132
20	medio_flor	CG12-613015	CP82-2043	CP81-1384/SP79-2233/CP33-224
21	medio_flor	CG12-400001	CG96-01	CP72-2086
22	medio_flor	CG12-324053	CG96-01	B7306
23	medio_flor	CG12-048041	L80-3	CP78-1610
24	medio_flor	CG12-030018	CP02-1143	CP78-1610
25	medio_flor	CG12-629043	CG96-01	CP65-357
26	medio_flor	CG12-333001	CP72-1210	CC85-92
27	medio_flor	CG12-629054	CG96-01	CP65-357
28	medio_flor	CG12-035011	CP02-1143	CP78-1610
29	medio_flor	CG12-076002	CP98-1056	RB726/CP88-1508
30	medio_flor	CG12-333045	CP72-1210	CC85-92
31	medio_flor	CG12-324058	CG96-01	B7306
32	medio_flor	CG12-135022	CG00-102	CP02-1143/CG98-46
33	medio_flor	CG12-036092	CP02-1143	CP81-1384
34	medio_flor	CG12-627080	CP73-1547	SP79-2233
35	medio_flor	CG12-324041	CG96-01	B7306
36	medio_flor	CG12-324021	CG96-01	B7306

37	medio_flor	CG12-326038	CG96-01	B74132
38	medio_flor	CG12-040033	CP29-103	/HoCP01-523CP77-1055
39	medio_flor	CG12-606035	CC85-63	CC84-75/CB38-22/B4362/Akoki
40	medio_flor	CG12-036096	CP02-1143	CP81-1384
41	medio_flor	CG12-610036	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
42	medio_flor	CG12-092002	CG97-97	CP02-1668
43	medio_flor	CG12-360006	CP72-1210	B74132/CP89-2143
44	medio_flor	CG12-365001	CP65-357	CC85-92
45	medio_flor	CG12-032010	CP02-1143	CP27-139
46	medio_flor	CG12-567003	CP89-2143	CG96-01
47	medio_flor	CG12-523006	CP88-1508	CP65-357/CG98-46
48	medio_flor	CG12-285001	CB38-22	CP82-2043/CP70-1133
49	medio_flor	CG12-326027	CG96-01	B74132
50	medio_flor	CG12-036004	CP02-1143	CP81-1384
51	medio_flor	CG12-638011	10P20602	BR9423
52	medio_flor	CG12-324003	CG96-01	B7306
53	medio_flor	CG12-194003	CP62-374	Gloria58
54	medio_flor	CG12-035030	CP02-1143	CP78-1610
55	medio_flor	CG12-629062	CG96-01	CP65-357
56	medio_flor	CG12-326032	CG96-01	B74132
57	medio_flor	CG12-318018	CB38-22	CG98-41/CC85-63
58	medio_flor	CG12-051114	CP75-1091	CP77-1055
59	medio_flor	CG12-637005	10P20601	WI9101
60	medio_flor	CG12-035024	CP02-1143	CP78-1610
61	medio_flor	CG12-260002	SP79-2233	CP73-1547/SP70-1284
62	medio_flor	CG12-005023	L56-20	CP81-1425
63	medio_flor	CG12-179001	CP82-2043	CP72-1210/CC85-92
64	medio_flor	CG12-534019	CP89-1431	CP63-588
65	medio_flor	CP72-2086	CP62-374	CP63-588
66	medio_flor	CG12-030005	CP02-1143	CP78-1610
67	medio_flor	CG12-338053	CP63-588	CP89-1431
68	medio_flor	CG12-041032	L80-3	CP78-1610
69	medio_flor	CG12-629039	CG96-01	CP65-357
70	medio_flor	CG12-386030	CP65-357	RB732908
71	medio_flor	CG12-616028	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
72	medio_flor	CG12-335003	CP73-1547	CP89-2143
73	medio_flor	CG12-338007	CP63-588	CP89-1431
74	medio_flor	CG12-338029	CP63-588	CP89-1431
75	medio_flor	CG12-022002	TCP83-3211	CP90-1222
76	medio_flor	CG12-008011	CP62-374	CP81-1425
77	medio_flor	CG12-621023	CP88-1508	L80-2/CP57-603
78	medio_flor	CG12-627025	CP73-1547	SP79-2233
79	medio_flor	CG12-627056	CP73-1547	SP79-2233
80	medio_flor	CG12-627090	CP73-1547	SP79-2233

81	medio_flor	CG12-620047	IJ76-319	Hinahina/IJ76-441/CP91-2249
82	medio_flor	CG12-584036	Mex57-29	CP65-357/CP88-1508
83	medio_flor	CG12-324191	CG96-01	B7306
84	medio_flor	CG12-324070	CG96-01	B7306
85	medio_flor	CG12-370002	CP63-588	CC85-63/CP89-2143
86	medio_flor	CG12-616012	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
87	medio_flor	CG12-040023	CP29-103	/HoCP01-523CP77-1055
88	medio_flor	CG12-611002	CP70-1133	CP68-1067CP70-321/CP72-1312
89	medio_flor	CG12-621017	CP88-1508	L80-2/CP57-603
90	medio_flor	CG12-135046	CG00-102	CP02-1143/CG98-46
91	medio_flor	CG12-135033	CG00-102	CP02-1143/CG98-46
92	medio_flor	CG12-134005	CG00-102	CP88-1165/CP73-1547
93	medio_flor	CG12-100004	Ja55-488	CP96-1329/Mex58-418
94	medio_flor	CG12-620020	IJ76-319	Hinahina/IJ76-441/CP91-2249
95	medio_flor	CG12-333050	CP72-1210	CC85-92
96	medio_flor	CG12-325017	CG96-01	B74132
97	medio_flor	CG12-626020	SP79-2233	CP73-1547/SP70-1284
98	medio_flor	CG12-014081	CP29-103	Akoki
99	medio_flor	CG98-78	Tuc68-19	CP57-603
100	medio_flor	CG12-253007	SP71-6180	CP70-1133
101	medio_flor	CG12-014085	CP29-103	Akoki
102	medio_flor	CG12-134007	CG00-102	CP88-1165/CP73-1547
103	medio_flor	CG12-092002	CG97-97	CP02-1668
104	medio_flor	CG12-040010	CP29-103	/HoCP01-523CP77-1055
105	medio_flor	CG12-610006	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
106	medio_flor	CG12-610018	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
107	medio_flor	CG12-616011	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
108	medio_flor	CG12-523058	CP88-1508	CP65-357/CG98-46
109	medio_flor	CG12-333033	CP72-1210	CC85-92
110	medio_flor	CG12-135014	CG00-102	CP02-1143/CG98-46
111	medio_flor	CG12-330081	Eros	CP75-1082
112	medio_flor	CG12-620010	IJ76-319	Hinahina/IJ76-441/CP91-2249
113	medio_flor	CG12-377002	CP63-588	CC85-92
114	medio_flor	CG12-032099	CP02-1143	CP27-139
115	medio_flor	CG12-325002	CG96-01	B74132
116	medio_flor	CG12-339026	CP63-588	Q96
117	medio_flor	CG12-534030	CP89-1431	CP63-588
118	medio_flor	CG12-135018	CG00-102	CP02-1143/CG98-46
119	medio_flor	CG12-338138	CP63-588	CP89-1431
120	medio_flor	CG12-534068	CP89-1431	CP63-588
121	medio_flor	CG12-059020	CP99-1944	CP03-2420
122	medio_flor	CG12-059094	CP99-1944	CP03-2420
123	medio_flor	CG12-312004	CG96-01	CP70-1133
124	medio_flor	CG12-326001	CG96-01	B74132

125	medio_flor	CG12-610029	CP65-357	CP63-588/CP68-1026/CP73-1547/CP72-370
126	medio_flor	CG12-523028	CP88-1508	CP65-357/CG98-46
127	medio_flor	CG12-626009	SP79-2233	CP73-1547/SP70-1284
128	medio_flor	CG12-059019	CP99-1944	CP03-2420
129	medio_flor	CG12-534010	CP89-1431	CP63-588
130	medio_flor	CG12-035119	CP02-1143	CP78-1610
131	medio_flor	CG12-338044	CP63-588	CP89-1431
132	medio_flor	CG12-337006	CP63-588	CP89-1431
133	medio_flor	CG12-239002	CP73-1547	CP81-1384
134	medio_flor	CG12-333019	CP72-1210	CC85-92
135	medio_flor	CG12-059015	CP99-1944	CP03-2420
1	medio_no_flor	CG12-122	RB739331	CP02-2188
2	medio_no_flor	CG12-317020	CB38-22	CG98-41/CC85-63
3	medio_no_flor	CG12-120	CP70-1133	CP01-2409/CP93-1361
4	medio_no_flor	CG12-338004	CP63-588	CP89-1431
5	medio_no_flor	CG12-338077	CP63-588	CP89-1431
6	medio_no_flor	CG12-548002	CC85-92	CP63-588
7	medio_no_flor	CG12-534037	CP89-1431	CP63-588
8	medio_no_flor	CG12-326045	CG96-01	B74132
9	medio_no_flor	CG12-135041	CG00-102	CP02-1143/CG98-46
10	medio_no_flor	CG12-523041	CP88-1508	CP65-357/CG98-46
11	medio_no_flor	CG12-324185	CG96-01	CP70-1133
12	medio_no_flor	CG12-311009	CG96-01	CP70-1133
13	medio_no_flor	CG12-376009	CP63-588	CC85-92
14	medio_no_flor	CG12-616003	CP91-2095	SP79-2233/CP91-2139/LM2002/Ja64-20
15	medio_no_flor	CG12-311028	CG96-01	CP70-1133
16	medio_no_flor	CG12-526009	CG96-01	CP70-1133
17	medio_no_flor	CG12-627037	CP73-1547	SP79-2233
18	medio_no_flor	CG12-324233	CG96-01	B7306
19	medio_no_flor	CG12-040026	CP29-103	CP77-1055/HoCP01-523
20	medio_no_flor	CG12-129	CP89-1431	CP63-588
21	medio_no_flor	CG12-008054	CP62-374	CP81-1425
22	medio_no_flor	CG12-621004	CP88-1508	L80-2/CP57-603
23	medio_no_flor	CG12-534006	CP89-1431	CP63-588
24	medio_no_flor	CG12-332013	CP72-1210	CP82-2043
25	medio_no_flor	CG12-057019	CP99-1707	CP47-193
26	medio_no_flor	CG12-606012	CC85-63	CC84-75/CB38-22/B4362/Akoki
27	medio_no_flor	CG12-050001	CP70-1133	CP91-1696
28	medio_no_flor	CG12-606103	CC85-63	CC84-75/CB38-22/B4362/Akoki
29	medio_no_flor	CG12-370003	CP63-588	CC85-63/CP89-2143
30	medio_no_flor	CG12-338073	CP63-588	CP89-1431
31	medio_no_flor	CG12-338098	CP63-588	CP89-1431
32	medio_no_flor	CG12-332015	CP72-1210	CP82-2043
33	medio_no_flor	CG12-617005	CP91-2139	CP91-2095/CP91-2161/Ja64-19/IJ76-521

34	medio_no_flor	CG12-031043	CP70-1133	CP01-2409/CP93-1361
35	medio_no_flor	CG12-339105	CP63-588	Q96
36	medio_no_flor	CG12-627035	CP73-1547	SP79-2233
37	medio_no_flor	CG12-318012	CB38-22	CG98-41/CC85-63
38	medio_no_flor	CG12-034098	CP02-1143	CP93-1361/CP78-1610
39	medio_no_flor	CG12-360009	CP72-1210	B74132/CP89-2143
40	medio_no_flor	CG12-131	CC85-63	CC84-75/CB38-22/B4362/Akoki
41	medio_no_flor	CG12-135	10#30010	B85342
42	medio_no_flor	CG12-512013	CG98-41	CB38-22/CC85-63
43	medio_no_flor	CG12-526022	CP75-1082	Eros
44	medio_no_flor	CG12-564005	CC85-92	CP65-357
45	medio_no_flor	CG12-339126	CP63-588	Q96
46	medio_no_flor	CG12-568016	CC85-63	CP63-588/CP89-2143
47	medio_no_flor	CG12-125	Eros	CP75-1082
48	medio_no_flor	CG12-386003	CP65-357	RB732908
49	medio_no_flor	CG12-338053	CP63-588	CP89-1431
50	medio_no_flor	CG12-337009	CP63-588	CP89-1431
51	medio_no_flor	CG12-633001	10#20014	B0209
52	medio_no_flor	CG12-229005	CP92-1401	CP72-1312
53	medio_no_flor	CG12-606100	CC85-63	CC84-75/CB38-22/B4362/Akoki
54	medio_no_flor	CG12-636002	10P20597	5
55	medio_no_flor	CG12-MSING#2	?	?
56	medio_no_flor	CG12-338087	CP63-588	CP89-1431
57	medio_no_flor	CG12-130	CP89-2143	SP91-2074
58	medio_no_flor	CG12-610024	CP65-357	CP63-588/CP68-1026/CP73-D5191547/CP72-370
59	medio_no_flor	CG12-041052	L80-3	CP78-1610
60	medio_no_flor	CG12-339073	CP63-588	Q96
61	medio_no_flor	CG12-629011	CG96-01	CP65-357
62	medio_no_flor	CG12-324187	CG96-01	B7306
63	medio_no_flor	CG12-034098	CP02-1143	CP93-1361/CP78-1610
64	medio_no_flor	CG12-626007	SP79-2233	CP73-1547/SP70-1284
65	medio_no_flor	CP88-1165	CL61-620	CP81-1302
66	medio_no_flor	CG12-330012	Eros	CP75-1082
67	medio_no_flor	CG12-133	CC85-63	CC84-75/CB38-22/B4362/Akoki
68	medio_no_flor	CG12-132	CC85-63	CC84-75/CB38-22/B4362/Akoki
69	medio_no_flor	CG12-136012	CP91-1696	CP70-1133
70	medio_no_flor	CG12-638010	10P20602	BR9423
71	medio_no_flor	CG12-621015	CP88-1508	L80-2/CP57-603
72	medio_no_flor	CG12-512009	CG98-41	CB38-22/CC85-63
73	medio_no_flor	CG12-533002	CP89-1431	CP63-588
74	medio_no_flor	CG98-78	Tuc68-19	CP57-603
75	medio_no_flor	CG12-338009	CP63-588	CP89-1431
76	medio_no_flor	CG12-128	CP88-1508	Mex79-431
77	medio_no_flor	CG12-123	CP91-1696	CP70-1133

78	medio_no_flor	CG12-531020	CP89-2143	CP73-1547
79	medio_no_flor	CG12-606042	CC85-63	CC84-75/CB38-22/B4362/Akoki
80	medio_no_flor	CG12-MSING#4	?	?
81	medio_no_flor	CG12-626022	SP79-2233	CP73-1547/SP70-1284
82	medio_no_flor	CG12-627077	CP73-1547	SP79-2233
83	medio_no_flor	CG12-317044	CB38-22	CG98-41/CC85-63
84	medio_no_flor	CG12-637001	10P20601	WI9101
85	medio_no_flor	CG12-627037	CP73-1547	SP79-2233
86	medio_no_flor	CG12-124	CG96-01	B7306
87	medio_no_flor	CP72-2086	CP62-374	CP63-588
88	medio_no_flor	CG12-620005	IJ76-319	Hinahina/IJ76-441/CP91-2249
89	medio_no_flor	CG12-564013	CC85-92	CP65-357
90	medio_no_flor	CG12-633001	10#20014	B0209
91	medio_no_flor	CG12-606095	CC85-63	CC84-75/CB38-22/B4362/Akoki
92	medio_no_flor	CG12-317043	CB38-22	CG98-41/CC85-63
93	medio_no_flor	CG12-337002	CP63-588	CP89-1431
94	medio_no_flor	CG12-127	CP89-2143	CP73-1547
95	medio_no_flor	CG12-526013	CP75-1082	Eros
96	medio_no_flor	CG12-629019	CG96-01	CP65-357
97	medio_no_flor	CG12-333031	CP72-1210	CC85-92
98	medio_no_flor	CG12-339111	CP63-588	Q96
99	medio_no_flor	CG12-317045	CB38-22	CG98-41/CC85-63
100	medio_no_flor	CG12-568029	CC85-63	CP63-588/CP89-2143
101	medio_no_flor	CG12-627060	CP73-1547	SP79-2233
102	medio_no_flor	CG12-136019	CP91-1696	CP70-1133
103	medio_no_flor	CG12-338008	CP63-588	CP89-1431
104	medio_no_flor	CG12-570001	CP89-2143	PR76-3385/CP72-2086
105	medio_no_flor	CG12-366001	CP65-357	CC85-92
106	medio_no_flor	CG12-564002	CC85-92	CP65-357
107	medio_no_flor	CG12-325038	CG96-01	B74132
108	medio_no_flor	CG12-051049	CP75-1091	CP77-1055
109	medio_no_flor	CP73-1547	CP66-1043	CP56-63
110	medio_no_flor	CG12-532009	CP88-1508	Mex79-431
111	medio_no_flor	CG12-011034	CP02-1143	CP77-1055
112	medio_no_flor	CG12-324147	CG96-01	B7306
113	medio_no_flor	CG12-320001	CG96-01	Q96/DB5/55
114	medio_no_flor	CG12-606030	CC85-63	CC84-75/CB38-22/B4362/Akoki
115	medio_no_flor	CG12-633003	10#20014	B0209
116	medio_no_flor	CG12-606011	CC85-63	CC84-75/CB38-22/B4362/Akoki
117	medio_no_flor	CG12-318022	CB38-22	CG98-41/CC85-63
118	medio_no_flor	CG12-512003	CG98-41	CB38-22/CC85-63
119	medio_no_flor	CG12-606102	CC85-63	CC84-75/CB38-22/B4362/Akoki
120	medio_no_flor	CG12-338007	CP63-588	CP89-1431
121	medio_no_flor	CG12-324084	CG96-01	B7306

122	medio_no_flor	CG12-627035	CP73-1547	SP79-2233
123	medio_no_flor	CG12-333077	CP72-1210	CC85-92
124	medio_no_flor	CG12-237004	CP73-1547	CP70-1133
125	medio_no_flor	CG12-638001	10P20602	BR9423
126	medio_no_flor	CG12-629049	CG96-01	CP65-357
127	medio_no_flor	CG12-324156	CG96-01	B7306
128	medio_no_flor	CG12-633002	10#20014	B0209
129	medio_no_flor	CG12-531012	CP89-2143	CP73-1547
130	medio_no_flor	CG12-330036	Eros	CP75-1082
131	medio_no_flor	CG12-317006	CB38-22	CG98-41/CC85-63
132	medio_no_flor	CG12-660001	?	?
133	medio_no_flor	CG12-318029	CB38-22	CG98-41/CC85-63
134	medio_no_flor	CG12-126	CP88-1508	CP65-357/CG98-46
135	medio_no_flor	CG12-627027	CP73-1547	SP79-2233
136	medio_no_flor	CG12-039002	CG98-36	HoCP01-523
137	medio_no_flor	CG12-134	10P20410	WI93904
138	medio_no_flor	CG12-121	RB739331	CP02-2188

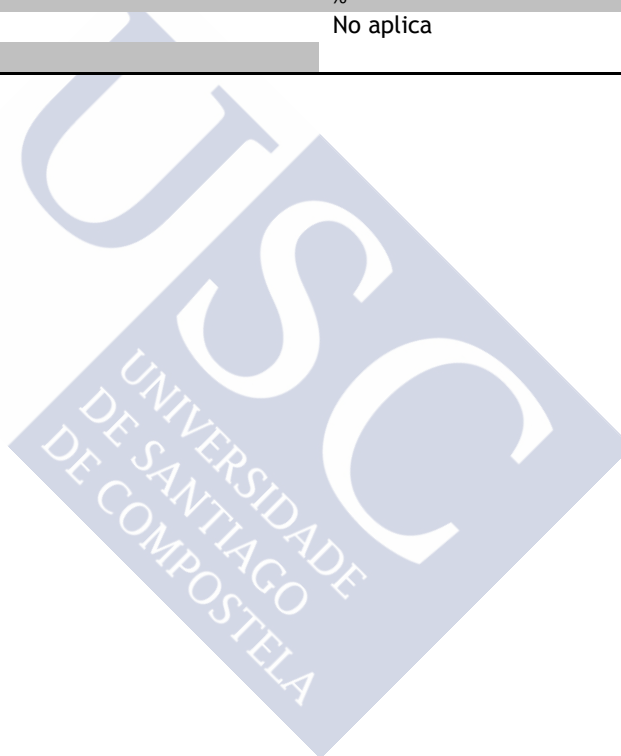
? = progenitor desconocido



ANEXO 2. LISTADO DE LAS MEDICIONES REALIZADAS EN CADA UNO DE LOS CONJUNTOS DE DATOS.

No.	Variable	Dimensional
1	Población	escala 1-9
2	Altura	escala 1-9
3	Diámetro	escala 1-9
4	Aspecto de Planta	escala 1-9
5	°Brix Tallo 1	%
6	°Brix Tallo 2	%
7	°Brix Tallo 3	%
8	Carbón	escala 0-10 ≤ 1
9	Escaldadura	escala (0-10) ≤ 1
10	Mosaico	escala (0-10) ≤ 2 B
11	Roya Marrón Incidencia	% (1-50) ≤ 15
12	Roya Marrón Reacción	escala (0-9) ≤ 5
13	Roya Naranja Incidencia	% (1-50) ≤ 10
14	Roya Naranja Reacción	escala (0-9) ≤ 5
15	Amarillamiento	escala (0-10) ≤ 5 B
16	SURCO 1 Total Tallos	frecuencia
17	SURCO 1 Total Tallos Con Flor	frecuencia
18	SURCO 2 Total Tallos	frecuencia
19	SURCO 2 Total Tallos Con Flor	frecuencia
20	SURCO 3 Total Tallos	frecuencia
21	SURCO 3 Total Tallos Con Flor	frecuencia
22	SURCO 4 Total Tallos	frecuencia
23	SURCO 4 Total Tallos Con Flor	frecuencia
24	ALTURA (m) Tallo 1	m
25	ALTURA (m) Tallo 2	m
26	ALTURA (m) Tallo 3	m
27	Diámetro (mm) Tallo 1	mm
28	Diámetro (mm) Tallo 2	mm
29	Diámetro (mm) Tallo 3	mm
30	Peso de 5 Tallos Molederos en kg. (PREVIO A	kg
31	Entrenudos Totales Tallo 1	frecuencia
32	Entrenudo con Corcho Tallo 1	frecuencia
33	Entrenudos Totales Tallo 2	frecuencia
34	Entrenudo con Corcho Tallo 2	frecuencia
35	Entrenudos Totales Tallo 3	frecuencia
36	Entrenudo con Corcho Tallo 3	frecuencia
37	Entrenudos Totales Tallo 4	frecuencia
38	Entrenudo con Corcho Tallo 4	frecuencia
39	Entrenudos Totales Tallo 5	frecuencia
40	Entrenudo con Corcho Tallo 5	frecuencia
41	Entrenudos Oquedad Leve Tallo 1	frecuencia
42	Entrenudos Oquedad Leve Tallo 2	frecuencia
43	Entrenudos Oquedad Leve Tallo 3	frecuencia
44	Entrenudos Oquedad Leve Tallo 4	frecuencia
45	Entrenudos Oquedad Leve Tallo 5	frecuencia
46	Entrenudos Oquedad Intermedia Tallo 1	frecuencia
47	Entrenudos Oquedad Intermedia Tallo 2	frecuencia
48	Entrenudos Oquedad Intermedia Tallo 3	frecuencia
49	Entrenudos Oquedad Intermedia Tallo 4	frecuencia
50	Entrenudos Oquedad Intermedia Tallo 5	Frecuencia
51	Entrenudos Oquedad Severa Tallo 1	Frecuencia
52	Entrenudos Oquedad Severa Tallo 2	Frecuencia
53	Entrenudos Oquedad Severa Tallo 3	Frecuencia
54	Entrenudos Oquedad Severa Tallo 4	Frecuencia
55	Entrenudos Oquedad Severa Tallo 5	Frecuencia

56	Oquedad Leve Cálculo	%
57	Oquedad Intermedia Cálculo	%
58	Oquedad Severa Calculo	%
59	TAH	t de Azúcar por ha (Cálculo)
60	Pol % Caña	%
61	TCH	t Caña por ha (Cálculo)
62	°Brix Jugo	%
63	Pureza Jugo	%
64	Contenido de Jugo	%
65	Población Total de Tallos	Frecuencia
66	Promedio Altura por Tallo	m (Cálculo del promedio)
67	Promedio Diámetro por Tallo mm	mm (Cálculo del promedio)
68	Promedio Peso por Tallo	Kg
69	Peso calculado Parcela	Kg
70	Total Tallos Con Flor	Frecuencia
71	Floración	% (Cálculo)
72	Corcho	% (Cálculo)
73	Fibra	%
74	Observaciones de Campo	No aplica



ANEXO 3. ÍNDICES DE SELECCIÓN PARA TODOS LOS GENOTIPOS (ENTRY) EN SU RESPECTIVO CONJUNTO DE DATOS (TRIAL).

Conjunto de datos	Genotipo	Índice									Selección (base de comparación)
		TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3	
litoral_flor_1a_soca	CG12-015002	23.44	-0.351	0.524	0.438	0.524	-0.469	-3.766	1.239	0.811	SI
litoral_flor_1a_soca	CG12-032004	21.14	-2.101	-1.200	-1.141	-1.200	0.891	0.799	1.419	-1.386	SI
litoral_flor_1a_soca	CG12-036004	21.43	-1.408	-1.727	-1.653	-1.727	1.165	1.488	1.223	-1.237	SI
litoral_flor_1a_soca	CG12-036096	26.53	2.255	1.688	1.690	1.688	-1.375	0.760	-1.558	2.394	SI
litoral_flor_1a_soca	CG12-051044	18.34	-4.514	-1.187	-1.157	-1.187	1.261	0.982	0.873	-0.342	NO
litoral_flor_1a_soca	CG12-065014	31.78	6.238	2.231	2.128	2.231	-1.404	-0.013	-1.959	0.137	SI
litoral_flor_1a_soca	CG12-105001	22.15	-1.505	-0.249	-0.201	-0.249	0.404	3.227	0.813	1.431	SI
litoral_flor_1a_soca	CG12-135002	23.40	-0.009	-0.873	-0.820	-0.873	0.754	1.312	1.020	-0.146	SI
litoral_flor_1a_soca	CG12-318023	36.14	10.195	3.121	3.049	3.121	-2.282	-1.735	-3.298	0.022	SI
litoral_flor_1a_soca	CG12-324175	21.56	-1.852	-0.446	-0.362	-0.446	0.486	2.267	0.511	-0.577	SI
litoral_flor_1a_soca	CG12-324199	22.59	-0.903	1.208	1.251	1.208	-1.054	1.652	0.648	1.528	SI
litoral_flor_1a_soca	CG12-326012	14.83	-6.861	0.143	0.117	0.143	-0.313	-1.836	0.704	1.195	NO
litoral_flor_1a_soca	CG12-326016	23.00	-0.658	-0.208	-0.273	-0.208	0.104	-1.612	0.664	-0.309	SI
litoral_flor_1a_soca	CG12-326036	23.60	0.382	1.346	1.352	1.346	-0.819	1.391	0.092	1.090	SI
litoral_flor_1a_soca	CG12-333034	34.57	8.981	2.140	2.111	2.140	-1.394	0.909	-2.205	1.011	SI
litoral_flor_1a_soca	CG12-333064	22.31	-1.228	-0.365	-0.387	-0.365	-0.007	-1.242	1.649	-0.987	SI
litoral_flor_1a_soca	CG12-338044	30.62	5.366	0.830	0.811	0.830	-0.469	-0.851	-2.766	-1.227	SI
litoral_flor_1a_soca	CG12-339002	20.55	-3.040	-1.210	-1.145	-1.210	0.965	0.551	0.525	-2.112	SI
litoral_flor_1a_soca	CG12-339013	19.63	-3.662	-0.838	-0.811	-0.838	0.556	-0.762	-0.281	-1.786	NO
litoral_flor_1a_soca	CG12-339042	30.53	5.668	0.854	0.822	0.854	-0.519	-0.017	-0.151	-0.123	SI
litoral_flor_1a_soca	CG12-339071	21.71	-2.078	-1.305	-1.339	-1.305	1.178	0.200	1.487	-0.932	NO
litoral_flor_1a_soca	CG12-339103	19.10	-3.782	-1.293	-1.327	-1.293	0.922	-2.606	-0.493	-0.397	NO
litoral_flor_1a_soca	CG12-533018	18.39	-4.345	0.195	0.247	0.195	-0.215	1.567	-0.592	-0.089	SI
litoral_flor_1a_soca	CG12-534044	28.12	3.348	-0.077	-0.089	-0.077	0.193	0.511	0.525	-0.548	SI
litoral_flor_1a_soca	CG12-534046	23.98	-0.206	-0.907	-0.831	-0.907	0.725	2.141	1.106	-0.786	SI
litoral_flor_1a_soca	CG12-596006	28.94	4.211	0.125	0.130	0.125	0.184	1.518	-1.335	-0.507	SI
litoral_flor_1a_soca	CG12-613003	22.63	-0.871	-1.144	-1.142	-1.144	0.998	0.157	0.499	-0.038	NO
litoral_flor_1a_soca	CG12-621002	18.02	-4.188	-0.623	-0.623	-0.623	0.214	-0.418	1.552	-0.100	NO
litoral_flor_1a_soca	CG12-621013	22.63	-1.016	-0.004	-0.063	-0.004	-0.224	-2.019	1.167	1.632	NO
litoral_flor_1a_soca	CG12-621037	16.82	-5.816	-0.423	-0.432	-0.423	-0.124	-3.290	-0.808	0.627	NO
litoral_flor_1a_soca	CG12-621060	18.56	-4.598	-0.404	-0.415	-0.404	-0.073	-2.351	1.110	-0.130	NO
litoral_flor_1a_soca	CG12-621065	18.90	-4.098	-0.129	-0.061	-0.129	-0.177	1.345	0.564	1.901	NO
litoral_flor_1a_soca	CG12-621067	32.76	7.724	-1.461	-1.458	-1.461	1.461	1.450	-0.533	-1.112	SI
litoral_flor_1a_soca	CG12-627080	18.86	-4.117	0.472	0.458	0.472	-0.401	0.424	0.017	0.912	NO
litoral_flor_1a_soca	CG12-629059	20.31	-2.347	0.037	-0.050	0.037	-0.245	-3.377	1.607	0.214	NO
litoral_flor_1a_soca	CG98-46	23.33	-0.395	0.406	0.439	0.406	-0.479	1.987	-1.368	1.248	SI
litoral_flor_1a_soca	CGB12-635007	34.48	8.687	0.119	0.096	0.119	0.054	-1.665	-1.612	-1.258	SI
litoral_flor_1a_soca	CGB12-635058	26.52	2.798	0.778	0.741	0.778	-0.789	0.088	0.068	0.456	SI
litoral_flor_1a_soca	CP72-2086	24.44	0.043	-0.455	-0.428	-0.455	0.468	0.556	-0.218	-0.627	SI
litoral_flor_1a_soca	CP73-1547	24.19	0.053	0.313	0.326	0.313	-0.150	0.274	-1.906	0.150	SI
litoral_flor_plantia	CG12-014008	18.17	-3.127	0.279	0.383	0.383	-0.102	1.454	1.071	1.884	NO
litoral_flor_plantia	CG12-032004	21.39	0.122	0.588	0.642	0.642	0.589	-1.877	0.384	1.535	NO
litoral_flor_plantia	CG12-036004	18.73	-0.959	-0.892	-0.963	-0.963	0.207	0.217	0.376	1.050	SI
litoral_flor_plantia	CG12-040023	21.81	-0.983	0.836	0.775	0.775	0.440	0.501	0.035	1.022	SI
litoral_flor_plantia	CG12-045002	20.24	-1.119	0.302	0.392	0.392	0.185	1.698	-0.573	0.936	SI
litoral_flor_plantia	CG12-116008	21.78	0.421	-0.394	-0.396	-0.396	-0.924	-1.549	-2.356	0.747	NO

Conjunto de datos	Genotipo	TAH	BLUP	ESIM	RESIM	Índice					Selección (base de comparación)
						KN	Smith	CP1	CP2	CP3	
litoral_flor_plantia	CG12-311014	28.79	2.956	-0.804	-0.782	-0.782	-0.096	0.069	-1.050	0.721	SI
litoral_flor_plantia	CG12-324082	20.54	-0.945	-1.159	-1.248	-1.248	0.098	0.595	0.220	0.721	NO
litoral_flor_plantia	CG12-324199	21.28	-0.717	1.494	1.627	1.627	1.056	0.597	0.032	0.648	SI
litoral_flor_plantia	CG12-333034	19.05	-1.162	2.087	1.977	1.977	-0.187	-0.796	0.212	0.584	NO
litoral_flor_plantia	CG12-338044	18.60	-2.006	1.916	1.818	1.818	0.580	0.285	1.241	0.575	NO
litoral_flor_plantia	CG12-339013	14.04	-3.165	-0.860	-0.926	-0.926	-0.084	-2.064	-0.226	0.518	NO
litoral_flor_plantia	CG12-339042	22.73	-0.442	0.443	0.256	0.256	-0.258	0.857	-0.278	0.348	SI
litoral_flor_plantia	CG12-339103	20.09	-0.574	0.197	0.055	0.055	-0.911	-0.197	1.182	0.247	NO
litoral_flor_plantia	CG12-360004	17.00	-4.000	-3.562	-3.537	-3.537	0.782	1.058	1.164	0.241	NO
litoral_flor_plantia	CG12-387001	24.12	0.787	-2.294	-2.291	-2.291	-1.108	1.704	0.153	0.136	SI
litoral_flor_plantia	CG12-508001	24.32	1.294	0.678	0.737	0.737	0.141	2.342	0.272	-0.007	SI
litoral_flor_plantia	CG12-533018	18.66	-1.389	0.277	0.346	0.346	0.405	-0.967	0.629	-0.096	NO
litoral_flor_plantia	CG12-534069	18.02	-0.820	1.313	1.441	1.441	-0.382	-2.343	0.232	-0.155	NO
litoral_flor_plantia	CG12-534083	30.52	5.502	0.171	0.238	0.238	-0.545	0.469	-1.417	-0.160	SI
litoral_flor_plantia	CG12-606034	12.75	-4.795	-0.486	-0.596	-0.596	0.706	-4.819	2.305	-0.231	NO
litoral_flor_plantia	CG12-613003	29.26	3.401	-1.241	-1.345	-1.345	-0.872	0.469	-1.954	-0.375	SI
litoral_flor_plantia	CG12-621044	22.25	-0.908	0.048	0.099	0.099	0.889	-2.221	-1.951	-0.424	SI
litoral_flor_plantia	CG12-621060	27.00	3.414	-1.331	-1.375	-1.375	0.059	-0.671	-1.932	-0.583	SI
litoral_flor_plantia	CG12-621064	16.82	-3.971	-0.073	0.049	0.049	0.233	-1.034	-0.514	-0.662	NO
litoral_flor_plantia	CG12-627080	32.33	6.080	1.354	1.401	1.401	-0.067	0.167	-0.988	-0.684	SI
litoral_flor_plantia	CG98-46	19.75	-1.410	1.979	2.045	2.045	-0.122	1.445	1.667	-1.291	NO
litoral_flor_plantia	CGB12-635007	25.94	2.805	0.706	0.591	0.591	0.172	2.448	0.008	-1.395	SI
litoral_flor_plantia	CGB12-635058	18.71	-2.228	-1.020	-0.965	-0.965	-0.034	-0.227	2.459	-1.647	NO
litoral_flor_plantia	CP72-2086	35.05	6.753	-1.645	-1.636	-1.636	-0.487	2.404	-0.501	-1.728	NO
litoral_flor_plantia	CP73-1547	24.82	1.184	1.092	1.186	1.186	-0.362	-0.013	0.095	-2.473	NO
litoral_no_flor_plantia	CG12-005031	20.77	-1.510	0.357	0.543	-0.020	-0.348	-1.646	1.304	0.045	SI
litoral_no_flor_plantia	CG12-018003	14.35	-6.556	-1.034	-1.287	0.025	0.362	2.096	0.390	0.514	NO
litoral_no_flor_plantia	CG12-030014	21.13	-1.382	1.084	1.174	-0.260	-0.028	1.098	0.505	-0.153	SI
litoral_no_flor_plantia	CG12-036037	16.75	-4.376	-0.328	-0.220	-0.382	1.775	-0.596	0.742	-0.514	SI
litoral_no_flor_plantia	CG12-060016	15.75	-4.729	-1.352	-1.679	-0.452	0.470	1.162	0.679	-0.337	SI
litoral_no_flor_plantia	CG12-101	23.42	0.282	-0.187	-0.250	0.252	0.474	-1.830	0.732	-0.292	SI
litoral_no_flor_plantia	CG12-102	19.11	-2.719	-1.027	-1.016	0.189	0.045	-0.304	1.670	0.041	SI
litoral_no_flor_plantia	CG12-103	14.32	-5.326	-0.243	-0.515	-0.164	-0.459	0.831	0.350	-1.312	NO
litoral_no_flor_plantia	CG12-104	25.09	1.123	-1.018	-1.163	-0.193	0.628	-1.723	1.079	-0.131	SI
litoral_no_flor_plantia	CG12-105	22.75	0.517	-0.534	-0.577	0.576	-0.790	1.510	-2.110	-3.269	SI
litoral_no_flor_plantia	CG12-106	11.72	-7.561	-1.815	-2.318	-0.169	0.544	1.215	2.415	0.020	NO
litoral_no_flor_plantia	CG12-106001	29.03	4.163	1.362	1.514	-0.027	-0.240	-1.455	-0.613	0.434	SI
litoral_no_flor_plantia	CG12-107	17.87	-3.237	-1.644	-1.705	-0.175	1.559	0.159	1.394	-0.274	NO
litoral_no_flor_plantia	CG12-108	22.37	-0.515	-0.906	-1.045	-0.099	0.511	0.685	1.094	0.582	SI
litoral_no_flor_plantia	CG12-109	27.30	2.542	1.188	1.049	0.571	-1.572	-0.682	-1.631	-0.427	SI
litoral_no_flor_plantia	CG12-110	22.04	-0.614	0.560	0.326	0.223	-0.359	2.009	1.379	1.838	SI
litoral_no_flor_plantia	CG12-111	36.78	9.206	-0.214	-0.085	-0.563	0.736	-2.139	-0.236	0.611	SI
litoral_no_flor_plantia	CG12-112	21.60	-0.966	-0.360	-0.498	0.082	-0.031	1.070	-0.128	0.069	SI
litoral_no_flor_plantia	CG12-113	24.12	0.858	0.013	0.118	-0.233	1.588	-2.686	0.610	-0.823	SI
litoral_no_flor_plantia	CG12-114	22.03	-0.946	0.755	0.478	0.621	-0.835	1.421	1.384	2.060	SI
litoral_no_flor_plantia	CG12-115	25.84	2.324	-1.682	-1.949	0.526	-1.127	2.234	-1.019	1.707	SI
litoral_no_flor_plantia	CG12-116	33.75	7.920	2.642	2.902	0.024	-0.503	-1.685	-0.620	0.310	SI
litoral_no_flor_plantia	CG12-276003	14.70	-5.774	-0.018	0.138	-0.257	-0.325	1.592	0.426	-0.509	NO
litoral_no_flor_plantia	CG12-311019	29.38	4.244	1.078	1.313	0.513	-1.201	-1.430	-1.261	0.731	SI

Conjunto de datos	Genotipo	Índice									Selección (base de comparación)
		TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3	
litoral_no_flor_plantia	CG12-311020	24.05	1.090	-0.422	-0.255	-0.120	0.197	-0.817	-2.301	-0.338	SI
litoral_no_flor_plantia	CG12-317012	20.35	-1.506	-0.437	-0.487	-0.224	0.815	0.745	-1.432	-1.987	SI
litoral_no_flor_plantia	CG12-317021	20.64	-1.294	-1.149	-0.615	-0.560	2.013	0.193	0.719	-1.753	SI
litoral_no_flor_plantia	CG12-318005	21.24	-1.198	-2.102	-2.378	0.168	0.215	1.376	1.282	0.257	SI
litoral_no_flor_plantia	CG12-324094	32.65	6.100	0.776	0.765	0.104	-1.154	-0.893	-2.616	-0.590	SI
litoral_no_flor_plantia	CG12-326030	28.34	4.381	1.018	0.915	0.321	-1.012	-0.549	1.083	1.088	SI
litoral_no_flor_plantia	CG12-330012	17.47	-3.110	0.857	0.914	-0.002	-0.032	-0.654	0.433	-1.551	SI
litoral_no_flor_plantia	CG12-339025	18.99	-2.163	0.265	0.126	0.319	-1.134	2.146	0.861	1.314	SI
litoral_no_flor_plantia	CG12-339052	21.30	-1.228	-1.210	-1.385	0.273	0.268	1.123	0.844	0.369	SI
litoral_no_flor_plantia	CG12-526031	22.75	-0.158	2.232	2.600	-0.317	-0.372	-2.934	-0.272	-0.587	SI
litoral_no_flor_plantia	CG12-532007	27.19	2.986	0.736	1.099	-0.400	0.540	-2.017	0.760	0.118	SI
litoral_no_flor_plantia	CG12-606129	20.01	-1.325	-0.468	-0.294	-0.168	0.166	2.926	-1.138	-1.498	SI
litoral_no_flor_plantia	CG12-620015	21.76	-0.856	-1.267	-1.488	-0.043	0.110	1.575	-1.676	1.127	SI
litoral_no_flor_plantia	CP72-2086	28.56	3.966	1.947	2.300	-0.103	0.201	-2.216	-0.996	0.343	SI
litoral_no_flor_plantia	CP88-1165	21.17	-1.631	-0.713	-0.680	-0.406	0.644	-1.359	-1.586	0.587	SI
litoral_no_flor_plantia	SP71-6161	29.17	4.699	2.451	2.501	0.454	-1.619	-0.783	0.044	1.117	SI
litoral_no_flor_plantia	SP79-1287	29.19	4.279	0.813	1.116	0.097	-0.721	1.229	-2.542	1.065	SI
medio_flor_1a_soca	CG12-005023	15.34	-0.111	0.453	0.527	0.405	0.155	1.378	0.728	-0.564	SI
medio_flor_1a_soca	CG12-030018	15.77	-0.132	0.443	0.512	0.490	0.188	-2.488	-0.779	-0.438	SI
medio_flor_1a_soca	CG12-035024	11.97	-1.677	-0.013	-0.024	0.137	-0.060	-1.164	2.554	-0.869	NO
medio_flor_1a_soca	CG12-036004	12.22	-1.878	0.011	0.012	0.125	-0.089	2.136	1.013	0.296	NO
medio_flor_1a_soca	CG12-036096	15.79	-0.097	1.179	1.102	1.440	1.756	0.655	-0.764	-1.264	NO
medio_flor_1a_soca	CG12-040023	15.49	-0.451	1.127	1.052	1.532	1.816	-2.507	1.423	0.837	NO
medio_flor_1a_soca	CG12-048041	20.44	2.562	-1.623	-1.550	-1.522	-1.394	1.193	-1.689	1.280	SI
medio_flor_1a_soca	CG12-059094	16.00	0.489	-1.628	-1.547	-1.511	-1.397	-0.196	-1.881	-1.248	SI
medio_flor_1a_soca	CG12-134007	17.75	0.784	0.372	0.428	0.230	0.234	-1.573	0.797	0.921	SI
medio_flor_1a_soca	CG12-194003	16.39	0.483	0.494	0.533	0.231	0.177	1.041	-0.649	0.101	SI
medio_flor_1a_soca	CG12-318018	18.60	1.715	1.343	1.147	1.588	1.726	-1.312	0.239	-0.070	SI
medio_flor_1a_soca	CG12-324003	13.99	-0.934	1.399	1.212	1.548	1.674	0.446	0.233	0.607	NO
medio_flor_1a_soca	CG12-325061	15.89	0.212	0.026	-0.040	-0.018	-0.089	-0.151	0.597	1.859	SI
medio_flor_1a_soca	CG12-326001	10.81	-2.795	-0.037	-0.087	0.053	-0.047	1.802	-0.454	-2.396	NO
medio_flor_1a_soca	CG12-330081	16.54	0.516	0.403	0.401	0.139	0.084	1.026	0.231	0.147	SI
medio_flor_1a_soca	CG12-333019	18.58	1.674	0.407	0.387	0.168	0.104	0.718	-0.921	1.334	SI
medio_flor_1a_soca	CG12-333050	20.46	2.094	-1.417	-1.298	-1.598	-1.392	0.500	-1.179	1.138	SI
medio_flor_1a_soca	CG12-338007	16.86	0.833	-1.408	-1.296	-1.600	-1.389	-2.938	-0.276	-0.343	SI
medio_flor_1a_soca	CG12-338053	12.04	-1.909	0.816	0.789	0.505	0.545	0.663	0.487	0.092	NO
medio_flor_1a_soca	CG12-377002	17.73	0.904	0.778	0.761	0.503	0.562	-0.747	-1.026	-0.432	SI
medio_flor_1a_soca	CG12-523006	19.87	2.262	-0.881	-0.830	-0.845	-0.692	-2.260	-1.663	0.063	SI
medio_flor_1a_soca	CG12-523009	13.84	-0.969	-0.904	-0.857	-0.804	-0.662	-1.618	-0.506	-0.405	SI
medio_flor_1a_soca	CG12-523028	17.78	1.177	0.192	0.210	-0.064	-0.091	0.818	0.279	1.158	SI
medio_flor_1a_soca	CG12-523058	12.51	-2.139	0.191	0.218	-0.082	-0.104	1.752	1.059	-0.598	NO
medio_flor_1a_soca	CG12-534090	17.50	1.211	-1.252	-1.439	-0.331	-0.452	-1.367	0.845	2.277	SI
medio_flor_1a_soca	CG12-610036	9.14	-3.703	-1.295	-1.474	-0.268	-0.417	1.415	0.911	-2.365	NO
medio_flor_1a_soca	CG12-611002	16.42	0.362	1.167	1.192	0.739	0.705	1.090	-0.635	-0.883	SI
medio_flor_1a_soca	CG12-621023	12.86	-1.400	1.156	1.197	0.698	0.682	-2.491	1.290	-1.603	NO
medio_flor_1a_soca	CG98-78	20.05	2.516	-0.746	-0.614	-0.977	-1.082	1.260	-0.557	1.315	SI
medio_flor_1a_soca	CP72-2086	12.62	-1.599	-0.751	-0.624	-0.910	-1.049	2.917	0.294	0.053	NO
medio_flor_plantia	CG12-005023	16.07	-0.899	1.507	1.511	-0.242	-0.365	0.745	0.378	-0.893	SI
medio_flor_plantia	CG12-030018	25.33	2.025	-0.024	-0.174	0.773	0.210	0.535	0.689	1.010	SI

Conjunto de datos	Genotipo	TAH	BLUP	ESIM	RESIM	Índice					Selección (base de comparación)
						KN	Smith	CP1	CP2	CP3	
medio_flor_plantia	CG12-032010	23.94	1.615	-1.349	-1.308	-0.809	-0.818	-2.740	0.245	2.003	NO
medio_flor_plantia	CG12-040023	18.54	-0.084	-1.386	-1.440	-0.198	-0.241	-0.896	1.206	-0.999	NO
medio_flor_plantia	CG12-092002	21.66	0.832	-0.928	-0.969	0.563	0.364	-1.079	-1.482	1.846	NO
medio_flor_plantia	CG12-135018	19.10	0.093	-1.022	-1.071	0.685	0.412	1.597	0.212	-1.500	NO
medio_flor_plantia	CG12-135022	14.77	-1.236	-1.925	-1.817	0.201	0.318	-0.536	-0.489	-1.806	NO
medio_flor_plantia	CG12-135033	14.60	-1.278	-2.759	-2.662	0.364	0.212	0.815	-0.323	0.249	NO
medio_flor_plantia	CG12-146003	13.83	-1.491	2.290	2.160	0.356	0.887	-2.077	-0.966	-1.218	NO
medio_flor_plantia	CG12-253007	24.66	1.929	-0.439	-0.577	1.367	0.896	2.344	1.721	0.988	NO
medio_flor_plantia	CG12-260002	23.07	1.299	-1.644	-1.437	-0.673	-1.034	0.278	0.051	0.717	SI
medio_flor_plantia	CG12-318018	20.17	0.456	0.890	1.023	-0.877	-0.810	-0.013	-0.014	0.720	SI
medio_flor_plantia	CG12-324003	15.52	-0.941	-2.224	-2.275	0.548	-0.266	-1.085	-1.301	0.596	SI
medio_flor_plantia	CG12-324070	14.95	-1.159	0.868	0.981	-1.398	-1.132	-3.315	-0.750	0.349	SI
medio_flor_plantia	CG12-324191	15.43	-1.018	2.296	2.409	-1.210	-0.996	1.107	0.082	1.791	NO
medio_flor_plantia	CG12-326001	16.45	-0.710	1.273	1.249	-0.238	0.260	-0.584	-3.255	1.603	SI
medio_flor_plantia	CG12-333017	17.21	-0.436	-1.171	-1.225	0.075	0.274	-2.104	-0.549	-1.028	NO
medio_flor_plantia	CG12-333050	21.94	1.005	-0.406	-0.312	-0.366	-0.416	0.735	0.668	-0.054	NO
medio_flor_plantia	CG12-335003	25.89	2.131	2.533	2.654	-1.051	0.027	-0.530	3.776	-0.251	SI
medio_flor_plantia	CG12-337006	11.40	-1.333	1.398	1.429	-0.521	-0.441	3.397	-1.247	0.038	NO
medio_flor_plantia	CG12-338007	19.20	0.110	0.139	0.082	0.173	-0.242	-2.289	0.670	-0.165	SI
medio_flor_plantia	CG12-523028	14.22	-1.347	-1.930	-1.983	0.492	-0.101	-0.432	-1.296	-1.551	SI
medio_flor_plantia	CG12-610018	11.63	-2.265	1.198	1.222	-0.054	-0.169	2.464	-2.074	-1.508	NO
medio_flor_plantia	CG12-620011	24.69	1.821	0.498	0.478	0.189	0.765	0.825	1.045	1.129	SI
medio_flor_plantia	CG12-621017	15.49	-1.037	0.097	-0.026	0.670	0.564	1.271	0.463	-1.203	SI
medio_flor_plantia	CG12-626020	11.97	-2.068	0.727	0.748	-0.412	-0.270	-0.216	-1.585	-1.448	NO
medio_flor_plantia	CG12-627080	17.88	-0.216	-0.643	-0.739	0.539	0.603	-0.890	0.427	-0.581	SI
medio_flor_plantia	CG98-78	25.05	2.013	-0.552	-0.714	1.119	1.056	0.810	1.550	0.165	SI
medio_flor_plantia	CP72-2086	23.54	1.492	1.513	1.474	0.407	0.806	2.356	0.057	-0.168	SI
medio_flor_plantia	CP73-1547	21.08	0.694	1.180	1.307	-0.477	-0.354	1.204	1.465	1.187	SI
medio_no_flor_1a_soca	CG12-008054	21.71	1.080	-0.571	-0.578	0.389	0.413	-1.270	-0.595	0.202	SI
medio_no_flor_1a_soca	CG12-034098	19.81	1.160	2.469	2.291	0.183	-0.169	-1.048	-0.561	0.389	SI
medio_no_flor_1a_soca	CG12-057019	17.04	-0.530	1.581	1.603	0.719	-0.279	0.177	-0.422	0.759	SI
medio_no_flor_1a_soca	CG12-121	15.44	-1.150	-1.581	-1.429	0.416	-0.083	1.241	0.949	-0.703	NO
medio_no_flor_1a_soca	CG12-122	20.48	0.720	-2.812	-2.487	0.106	-0.061	0.997	-0.305	0.083	SI
medio_no_flor_1a_soca	CG12-123	17.29	-0.503	-1.069	-0.797	0.419	-0.480	-1.226	1.175	-0.007	SI
medio_no_flor_1a_soca	CG12-124	16.95	-0.597	1.026	0.762	-0.013	-0.279	-0.980	0.438	-0.922	SI
medio_no_flor_1a_soca	CG12-125	19.49	0.158	0.472	0.255	-0.376	0.207	0.590	0.919	-1.073	SI
medio_no_flor_1a_soca	CG12-126	20.36	0.624	-0.039	0.082	-0.056	0.048	-0.562	-2.367	0.446	SI
medio_no_flor_1a_soca	CG12-130	24.50	1.870	-1.731	-1.426	-0.208	-0.127	3.457	-0.778	0.715	SI
medio_no_flor_1a_soca	CG12-132	23.53	1.628	-1.021	-0.980	0.304	0.696	-0.763	-1.385	0.763	SI
medio_no_flor_1a_soca	CG12-133	19.06	0.102	-0.999	-1.003	0.595	0.382	1.740	-0.132	0.231	SI
medio_no_flor_1a_soca	CG12-135	21.30	0.722	-0.945	-0.660	0.403	-0.041	0.108	0.754	0.818	SI
medio_no_flor_1a_soca	CG12-229005	31.25	4.123	0.541	0.778	0.451	-0.179	0.879	-0.936	2.330	SI
medio_no_flor_1a_soca	CG12-317020	15.71	-0.984	-0.877	-0.964	-0.097	0.111	1.795	-0.941	-1.979	NO
medio_no_flor_1a_soca	CG12-317043	15.65	-0.941	-1.987	-2.148	-0.456	0.349	0.181	-0.851	0.502	NO
medio_no_flor_1a_soca	CG12-317044	16.99	-0.490	0.923	0.989	0.143	-0.257	2.389	-0.892	0.970	NO
medio_no_flor_1a_soca	CG12-318012	22.11	1.128	1.316	1.515	0.242	-0.082	-1.399	-0.897	1.903	SI
medio_no_flor_1a_soca	CG12-318022	16.30	-0.757	1.165	1.237	0.040	-0.119	0.332	-0.822	-0.305	NO
medio_no_flor_1a_soca	CG12-324084	19.37	0.267	-1.283	-1.423	-0.517	-0.217	0.850	1.591	-0.277	SI
medio_no_flor_1a_soca	CG12-324147	20.84	0.791	-0.143	-0.199	0.084	-0.127	-1.058	3.658	1.505	NO

Conjunto de datos	Genotipo	Índice									Selección (base de comparación)
		TAH	BLUP	ESIM	RESIM	KN	Smith	CP1	CP2	CP3	
medio_no_flor_1a_soca	CG12-324156	18.41	-0.049	0.689	0.782	0.377	-0.059	0.260	-1.360	0.780	NO
medio_no_flor_1a_soca	CG12-324185	18.16	-0.091	-0.898	-0.757	-0.307	0.168	0.565	0.463	-0.767	SI
medio_no_flor_1a_soca	CG12-324187	20.26	0.473	0.768	0.908	-0.330	0.092	-1.880	-0.224	-0.110	SI
medio_no_flor_1a_soca	CG12-330036	20.39	0.540	2.399	2.350	-0.383	-0.318	-1.282	-0.016	-0.767	SI
medio_no_flor_1a_soca	CG12-332013	14.54	-1.389	0.905	0.551	-0.740	-0.243	-0.959	-0.304	-1.442	NO
medio_no_flor_1a_soca	CG12-332015	16.78	-0.635	1.213	0.772	-0.532	0.127	-1.407	0.264	-1.066	SI
medio_no_flor_1a_soca	CG12-333077	18.18	-0.161	2.657	2.508	0.116	-0.101	-1.395	0.952	0.760	SI
medio_no_flor_1a_soca	CG12-337009	16.27	-0.879	0.328	0.569	0.313	-0.438	0.559	-2.141	0.206	NO
medio_no_flor_1a_soca	CG12-360009	18.12	-0.219	0.305	0.448	0.214	-0.142	-1.078	0.035	-0.224	SI
medio_no_flor_1a_soca	CG12-512009	19.19	0.181	1.496	1.398	0.156	-0.113	-1.025	-0.784	0.592	SI
medio_no_flor_1a_soca	CG12-512013	20.26	0.479	3.629	3.515	-0.031	0.004	-2.296	0.254	0.433	SI
medio_no_flor_1a_soca	CG12-531020	16.69	-0.690	1.270	1.237	-0.059	0.351	0.767	0.626	-1.019	SI
medio_no_flor_1a_soca	CG12-533002	22.25	1.139	0.040	-0.017	0.025	0.244	0.055	0.631	0.154	SI
medio_no_flor_1a_soca	CG12-534006	16.01	-0.828	-1.183	-1.303	-0.468	-0.086	1.053	-0.645	-1.731	SI
medio_no_flor_1a_soca	CG12-548002	19.93	0.456	-1.078	-1.215	-0.868	0.145	0.718	-0.067	-0.131	SI
medio_no_flor_1a_soca	CG12-564013	13.75	-1.598	-1.365	-1.375	-0.042	0.167	1.703	-0.464	0.010	NO
medio_no_flor_1a_soca	CG12-570001	14.31	-1.381	-1.197	-1.127	-0.734	-0.278	-0.469	0.097	-1.046	NO
medio_no_flor_1a_soca	CG12-606012	13.02	-1.849	1.876	1.521	-0.630	-0.242	-2.443	0.164	-1.748	NO
medio_no_flor_1a_soca	CG12-606100	17.81	-0.345	-0.554	-0.647	-0.107	0.267	2.740	1.105	0.095	NO
medio_no_flor_1a_soca	CG12-606102	19.80	0.354	-2.607	-2.461	0.294	0.308	1.323	5.441	1.566	NO
medio_no_flor_1a_soca	CG12-626007	15.95	-0.909	-2.874	-3.153	-0.345	0.119	1.609	-0.234	-1.725	NO
medio_no_flor_1a_soca	CG12-627037	23.27	0.797	-0.607	-0.276	0.203	0.195	-1.413	0.304	0.861	SI
medio_no_flor_1a_soca	CG12-629011	18.64	-0.003	-0.157	-0.317	-0.366	0.204	0.950	-0.534	0.783	SI
medio_no_flor_1a_soca	CG12-637001	12.91	-1.887	-1.696	-1.433	0.722	0.380	-0.348	0.925	-0.990	NO
medio_no_flor_1a_soca	CG12-MSING#4	23.47	1.723	0.395	0.563	-0.020	-0.476	-1.144	-2.226	2.064	SI
medio_no_flor_1a_soca	CG98-78	16.09	-0.878	0.382	0.589	0.426	-0.434	-1.889	-0.895	-0.845	SI
medio_no_flor_1a_soca	CP72-2086	17.04	-0.505	-0.272	-0.504	0.463	0.444	-0.269	1.586	-1.416	SI
medio_no_flor_1a_soca	CP88-1165	17.86	-0.270	1.698	1.457	-0.117	0.007	0.569	-0.553	0.374	SI
medio_no_flor_plantia	CG12-031043	22.60	0.851	-0.542	-0.551	0.016	-0.332	-1.981	-0.574	1.242	SI
medio_no_flor_plantia	CG12-120	21.93	0.438	-4.051	-3.990	-0.367	-0.520	0.676	0.075	0.722	SI
medio_no_flor_plantia	CG12-121	21.02	0.636	-2.933	-2.862	-0.134	-0.228	0.907	0.890	-1.053	SI
medio_no_flor_plantia	CG12-122	16.95	-2.199	1.231	1.253	0.206	0.206	-0.980	0.967	0.271	NO
medio_no_flor_plantia	CG12-123	17.40	-1.679	1.257	1.242	-0.066	0.062	0.299	0.366	-0.178	SI
medio_no_flor_plantia	CG12-124	12.47	-4.153	1.407	1.404	-0.090	0.043	1.373	0.534	-0.926	NO
medio_no_flor_plantia	CG12-125	24.99	2.380	-0.026	-0.149	-0.060	-0.024	-1.231	1.157	0.388	SI
medio_no_flor_plantia	CG12-126	20.05	-0.122	-1.445	-1.698	-0.031	-0.044	1.165	0.917	-0.667	SI
medio_no_flor_plantia	CG12-127	20.39	-0.066	-0.369	-0.786	0.084	-0.092	-0.775	-0.859	-0.271	SI
medio_no_flor_plantia	CG12-128	24.76	1.898	1.048	0.733	0.187	-0.034	1.375	0.479	0.662	SI
medio_no_flor_plantia	CG12-129	31.59	5.455	1.654	1.664	0.140	0.097	0.502	0.063	1.374	SI
medio_no_flor_plantia	CG12-130	30.93	5.086	0.414	0.376	0.304	0.114	-1.487	0.211	1.949	SI
medio_no_flor_plantia	CG12-131	19.91	-0.592	-0.645	-0.560	0.083	0.139	-0.428	0.725	0.429	SI
medio_no_flor_plantia	CG12-132	20.53	0.051	-1.360	-1.054	-0.058	0.103	0.513	0.174	2.851	SI
medio_no_flor_plantia	CG12-133	17.38	-1.693	-0.145	0.110	0.103	0.283	-0.471	0.198	0.121	NO
medio_no_flor_plantia	CG12-134	16.92	-1.590	0.347	0.452	-0.048	0.062	-4.606	0.411	-1.520	NO
medio_no_flor_plantia	CG12-135	24.68	1.808	-0.809	-0.729	-0.037	-0.111	1.363	-0.453	-1.008	SI
medio_no_flor_plantia	CG12-317043	20.42	0.158	-0.098	-0.059	0.016	0.119	1.188	-1.473	-0.037	SI
medio_no_flor_plantia	CG12-324233	23.32	1.388	1.247	1.161	0.150	0.362	-0.707	-0.427	1.852	SI
medio_no_flor_plantia	CG12-512013	17.90	-1.444	0.440	0.558	-0.150	-0.116	1.129	0.222	-0.111	NO
medio_no_flor_plantia	CG12-548002	19.39	-0.732	0.703	0.620	-0.078	-0.274	0.418	-5.625	-1.151	SI

Conjunto de datos	Genotipo	TAH	BLUP	ESIM	RESIM	Índice					Selección (base de comparación)
						KN	Smith	CP1	CP2	CP3	
medio_no_flor_plantia	CG12-570001	17.49	-1.560	0.029	-0.214	-0.047	-0.064	-0.060	0.228	-1.675	SI
medio_no_flor_plantia	CG12-606011	16.12	-2.361	-1.441	-1.459	-0.224	-0.113	0.064	0.362	-1.432	NO
medio_no_flor_plantia	CG12-606095	19.88	-0.006	-1.306	-1.261	-0.297	-0.155	1.030	0.551	0.246	SI
medio_no_flor_plantia	CG12-606102	17.95	-1.259	-1.400	-1.206	-0.019	0.041	-1.345	0.296	-0.273	NO
medio_no_flor_plantia	CG12-627037	19.66	-0.565	0.807	0.636	0.152	0.262	-1.171	0.710	-1.627	SI
medio_no_flor_plantia	CG12-627037a	16.86	-1.994	2.652	2.309	0.067	0.079	-0.769	-1.169	-1.828	NO
medio_no_flor_plantia	CG98-78	27.92	3.512	1.807	2.099	0.113	0.206	1.511	-0.482	2.064	SI
medio_no_flor_plantia	CP72-2086	16.90	-1.646	1.528	1.961	0.084	-0.073	2.498	1.524	-0.417	NO



ANEXO 4. RESUMEN DEL ANÁLISIS DE VARIANZA PARA UN MODELO MIXTO, DONDE LA VARIABLE “VARIEDAD” (ENTRY) SE CONSIDERÓ FIJA, EL RESTO SE CONSIDERARON ALEATORIAS. VARIABLES DE RESPUESTA: SCORES DE 14 COMPONENTES PRINCIPALES (CPS).

Conjunto de Datos	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10	CP11	CP12	CP13	CP14
	7.410	4.140	1.987	3.963	5.734	2.834	2.963	1.713	2.139	2.814	2.303	1.584	1.791	0.832
	<0.0001	<0.0001	0.017	<0.0001	<0.0001	0.002	0.000	0.04841	0.020	0.001	0.005	0.112	0.050	0.717
	****	****	*	****	****	**	***	*	*	***	**	NS	NS	NS
litoral_flor_plantia	4.089	2.202	1.289	0.839	1.496	2.292	2.625	1.416	1.285	1.567	1.369	1.089	0.750	2.147
	0.001	0.016	0.246	0.674	0.163	0.019	0.005	0.243	0.280	0.110	0.232	0.407	0.784	0.019
	***	*	NS	NS	NS	*	**	NS	NS	NS	NS	NS	NS	*
litoral_no_flor_plantia	5.314	3.843	1.893	8.369	4.382	4.776	3.851	2.009	3.714	1.333	1.461	1.477	0.963	1.275
	<0.0001	<0.0001	0.05	<0.0001	<0.0001	<0.0001	<0.0001	0.06	<0.0001	0.18	0.12	0.15	0.55	0.26
	****	****	*	****	****	****	****	NS	****	NS	NS	NS	NS	NS
medio_flor_1a_soca	2.587	1.246	2.513	4.942	4.311	4.244	1.076	2.819	1.725	2.132	1.136	1.517	2.821	1.256
	0.021	0.276	0.049	<0.0001	<0.0001	0.000	0.442	0.015	0.162	0.023	0.367	0.261	0.003	0.272
	*	NS	*	****	****	***	NS	*	NS	*	NS	NS	**	NS
medio_flor_plantia	4.888	3.476	3.588	0.693	1.310	1.985	1.247	3.646	1.708	2.373	1.545	2.087	0.474	1.894
	0.025	0.001	0.042	0.803	0.236	0.065	0.325	0.007	0.078	0.012	0.180	0.185	0.975	0.217
	*	***	*	NS	NS	NS	NS	**	NS	*	NS	NS	NS	NS
medio_no_flor_1a_soca	2.056	5.255	3.680	3.571	3.885	1.714	1.452	1.467	1.648	2.696	0.909	1.653	1.149	1.285
	0.017	<0.0001	<0.0001	0.001	<0.0001	0.032	0.136	0.126	0.086	0.003	0.629	0.095	0.315	0.192
	*	****	****	***	****	*	NS	NS	NS	**	NS	NS	NS	NS
medio_no_flor_plantia	1.388	2.263	2.748	1.243	1.398	1.372	1.050	1.363	1.099	1.150	1.207	1.363	1.149	1.014
	0.075	0.000	<0.0001	0.169	0.071	0.083	0.410	0.103	0.354	0.279	0.203	0.120	0.269	0.469
	NS	***	****	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

